Spectral Properties of a Discrete Lévy-driven Metastable System

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Rat der
Fakultät für Mathematik und Informatik
der Friedrich-Schiller-Universität Jena

von Dipl.-Math. Toralf Burghoff,
geboren am 24.08.1986 in Jena

Gutachter:

1. Prof. Dr. Ilya Pavlyukevich, *Friedrich-Schiller-Universität Jena.*
2. Prof. Dr. Peter Imkeller, *Humboldt-Universität zu Berlin.*
3. Prof. Dr. Erika Hausenblas, *Montanuniversität Leoben.*

Tag der öffentlichen Verteidigung: 4. Dezember 2014

Danksagungen

Ich möchte mich aufrichtig bei meinem Doktorvater Prof. Dr. Ilya Pavlyukevich für die Betreuung während meiner Zeit als Doktorand bedanken. Er stand mir stets geduldig mit guten Ratschlägen zur Seite und gab wichtige Impulse, ohne die diese Arbeit nicht möglich gewesen wäre. Des Weiteren möchte ich einen großen Dank an meine Kollegen aussprechen, insbesondere an Dr. Philipp Rudolph, mit denen ich viele fruchtbare Diskussionen geführt habe. Auch meiner Familie und meinen Freunden gebührt ein Dankeschön für ihre Unterstützung während der Höhen und Tiefen der letzten vier Jahre.
Nicht zuletzt möchte ich dem Deutschen Akademischen Austauschdienst für die Bereitstellung eines Stipendiums danken, mit welchem es mir ermöglich wurde, ein wundervolles und produktives Jahr in Australien zu verbringen.

Acknowledgments

I want to thank my doctoral advisor Prof. Dr. Ilya Pavlyukevich sincerely for the mentoring during my time as a PhD student. He patiently provided good advices and important impules which were of great help to me and made this thesis possible. Furthermore, I would like to thank my dear colleagues, especially Dr. Philipp Rudolph, for many fruitful discussions. Also, a huge thanks goes to my family and friends who supported me during the ups and downs in the last four years.
Last bot not least I want to thank the German Academic Exchange Service for providing me with a scholarship that enabled me to spend a wonderful and productive year in Australia.

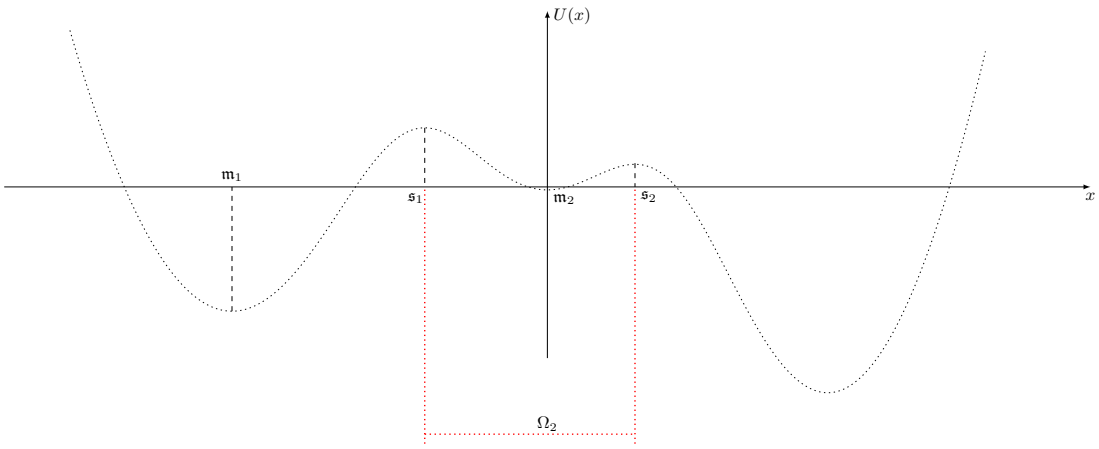
Contents

Einleitung und Zusammenfassung der Resultate

Wir betrachten als Ausgangspunkt das deterministische dynamische System

$$\dot{X}_t = -U'(X_t), \quad t \geq 0,$$

unter der Anfangsbedingung $X_0 = x$. Hierbei ist $U\colon \mathbb{R} \to \mathbb{R}$ ein Mehrtopfpotential mit lokalen Minima $\mathfrak{m}_1, \ldots, \mathfrak{m}_n,\ n \geq 1$, sowie $n-1$ lokalen Maxima $\mathfrak{s}_1, \ldots, \mathfrak{s}_{n-1}$. Die Potentialtöpfe Ω_i sind die Intervalle zwischen zwei zwei lokalen Maxima, d.h. $\Omega_i := (\mathfrak{s}_{i-1}, \mathfrak{s}_i)$.



Vom physikalischen Standpunkt her ist es klar, dass die Trajektorie $(X_t)_{t\geq 0}$ eines Teilchens im vom U erzeugten Kraftfeld, je nach Startpunkt, gegen ein lokales Minimum von U konvergiert bzw. in einem Extremum verbleibt. Ein Übergang zwischen den Potentialtöpfen ist nicht möglich.

Die Situation ändert sich, wenn man in obiger Gleichung eine kleine (stochastische) Störung

addiert. Man erhält dann die stochastische Differentialgleichung

$$X_t^\varepsilon = x - \int_0^t U'(X_s^\varepsilon)ds + \varepsilon\eta_t \tag{0.0.1}$$

mit einem stochastischen Prozess $\eta = (\eta_t)_{t\geq 0}$ als Rauschterm, dessen Intensität durch den Parameter $\varepsilon > 0$ kontrolliert wird. Fraglos am häufigsten betrachtet wurde der Gaußsche Fall mit $\eta = B$, wobei B eine Brownsche Bewegung ist. Dieser wird in der Literatur in der Regel mit dem Begriff *Diffusion* bezeichnet.[1]

Diffusionsgleichungen werden seit Jahrzehnten intensiv studiert und besitzen zahlreiche Anwendungen in Physik, Biologie, Finanzmathematik etc. Die Standardliteratur zum Thema stochastische Differentialgleichungen umfasst unter anderem die Bücher [Ok] und [KaSh] während die Hauptreferenz für Untersuchungen von Gleichung 0.0.1 im Grenzverhalten $\varepsilon \to 0$, auch bekannt als "small noise limit", wohl das Buch [FrWe] von Freidlin und Wentzell ist. Ein Nachteil bei der Modellierung mit dem stochastischen Prozess der Brownschen Bewegung, welche in diesem Kontext auch häufig als *Gaußsches weißes Rauschen* bezeichnet wird, ist, dass das Auftreten *extremer Ereignisse*, beispielsweise Börsencrashs oder Naturkatastrophen, durch die exponentiell schnell abfallenden Verteilungsschwänze der Normalverteilung nicht akkurat beschrieben werden kann. Aus diesem Grund wurden in den letzten Jahren Gleichungen der Form (0.0.1) mit Rauschtermen η betrachten, bei denen instantane Sprünge möglich sind. Eine wichtige solche Klasse von stochastischen Prozessen stellen die symmetrischen α-stabilen Lévy-Prozesse dar, wobei $0 < \alpha < 2$ der sogenannte Stabilitätsindex ist. Diese werden häufig als *Lévy-Rauschen* oder, im Zusammenhang mit der Modellierung von physikalischen Problemen, auch als *Lévy Flights* bezeichnet.

Ein konkretes Beispiel für die Modellierung eines geophysikalischen Phänomens durch Gleichung (0.0.1) mit einem Mehrtopfpotential U und einem solchen Lévy-Prozess wird in den Arbeiten von Ditlevsen ([Di1], [Di2]) diskutiert. Dabei untersuchte der Autor die Temperaturentwicklung in der nördlichen Hemisphäre in den letzten 100 000 Jahren und

[1] Als Anmerkung sei gesagt, dass man unter dem Begriff "Diffusion" natürlich auch allgemeinere Driftkomponenten als nur Mehrtopfpotentiale zulässt. Auch die externe Steuerung der Intensität durch eine Parameter ε wird häufig weggelassen.

stellte durch die Analyse von Eiskernmessungen fest, dass es in diesem Zeitraum etwa 25 abrupte Temperaturänderungen gab, die als solche "katastrophalen" Ereignisse interpretiert werden können. Ein einfaches Modell zur Beschreibung dieser Beobachtung ist daher die Verwendung eines eines Zweitopfpotentials, bei welchem die beiden lokalen Minima mit den Zuständen "kalt" und "warm" identifiziert werden können.

Wir betrachten nun ebenfalls Gleichung (0.0.1) mit einem symmetrischen α-stabilen Lévy-Rauschen. Die vorliegende Arbeit befasst sich mit dem Studium einer Markov-Kette, die durch eine Diskretisierung dieser Gleichung induziert wird.
Das erste Kapitel beginnt mit einem Überblick über die verwendeten Bezeichnungen und Definitionen. Danach werden die für die spätere Arbeit notwendigen theoretischen Grundlagen bereitgestellt. Wir starten dabei mit der Einführung von Laplace-Transformierten von Zufallsvariablen, wobei wir einen im Vergleich zur Standardliteratur geringfügig allgemeineren Zugang gewählt haben. Anschließend werden kurz die Definitionen und Ergebnisse aus der Theorie von Markov-Prozessen angegeben, welche in dieser Dissertation gebraucht werden. Als letztes gehen wir detaillierter auf symmetrische α-stabile Lévy-Prozesse ein.

Das zweite Kapitel dient dann der Einführung zweier zeitdiskreter Markov-Ketten Z^0 und Z^ε auf einem gemeinsamen endlichen Zustandsraum $\mathcal{S}$. Diese Ketten sollen eine Diskretisierung der Lösung von Gleichung (0.0.1) für die Fälle $\varepsilon = 0$ sowie $\varepsilon > 0$ darstellen. Als Ansatz dafür wählen wir für die Zeitachse ein Euler-Schema mit Schrittweite $h > 0$. Verwendet man zusätzlich die Selbstähnlichkeit von Lévy-Prozessen, so erhält man damit die Rekursionsgleichung

$$\tilde{X}_0 = x, \quad \tilde{X}^\varepsilon_{kh} = \tilde{X}^\varepsilon_{(k-1)h} - hU'(\tilde{X}^\varepsilon_{(k-1)h}) + h^{\frac{1}{\alpha}}\varepsilon\xi_k, \quad k \geq 1,$$

wobei $(\xi_k)_{k\geq 1}$ eine Folge unabhängig und identisch verteilter Zufallsvariablen mit $\xi_1 \stackrel{d}{=} L_1$ ist. Um den Zustandsraum $\mathcal{S}$ zu definieren, schränken wir unsere Betrachtung auf ein kompaktes Intervall $[-R, R]$ ein, welches alle lokalen Minima von U enthält. Anschließend wird dieses Intervall in N nicht notwendigerweise gleichgroße, disjunkte Teilintervalle eingeteilt,

wobei jedes eine maximale Länge $\delta > 0$ hat. Aus jedem dieser Teilintervalle wird jeweils ein Punkt als Zustand festgelegt, wobei wir zusätzlich fordern, dass die Minima von U ausgewählt und die Maxima ausgeschlossen werden müssen. Für einen Zustand $y \in \mathcal{S}$ bezeichne I_y dass Teilintervall, welches y enthält. Dann sind die Übergangswahrscheinlichkeiten $p^0_{x,y},\ x, y \in \mathcal{S}$, für Z^0 bzw. $p^\varepsilon_{x,y}$ für Z^ε definiert durch

$$p^0_{x,y} \coloneqq \mathbb{P}(x - hU'(x) \in I_y) \quad \text{bzw.} \quad p^\varepsilon_{x,y} \coloneqq \mathbb{P}(x - hU'(x) + h^{\frac{1}{\alpha}}\varepsilon\xi_1 \in I_y).$$

Durch Ausnutzen bekannter Reihenentwicklungen für die Verteilungsschwänze $\mathbb{P}(\xi_1 > a),\ a \to \infty$, können wir folgende Aussagen über die Übergangswahrscheinlichkeiten beweisen:

Proposition 2.2.1. *Es seien $\mathbf{P}^0 = (p^0_{x,y})_{x,y\in\mathcal{S}}$ und $\mathbf{P}^\varepsilon = (p^\varepsilon_{x,y})_{x,y\in\mathcal{S}}$ die Übergangsmatrizen für Z^0 bzw. Z^ε. Dann gilt:*

(i) Für jedes $x \in \mathcal{S}$ gibt es genau einen Zustand $y^* = y^*(x)$ mit $p^0_{x,y^*} = 1$.

(ii) Es sei $x \in \mathcal{S}$ und y^* der eindeutig bestimmte Zustand aus (i). Dann gilt $1 - p^\varepsilon_{x,y^*} = O(\varepsilon^\alpha)$ und $p^\varepsilon_{x,y} = O(\varepsilon^\alpha)$ für alle anderen Zustände $y \neq y^*$.

Im letzten Abschnitt untersuchen wir dann das asymptotische Verhalten der stationären Verteilung von Z^ε. Dafür müssen wir eine Größe aufgreifen, nämlich eine zeitstetige Markov-Kette Y mit Zustandsraum $\{\mathfrak{m}_1, \dots, \mathfrak{m}_n\}$, welche wir erst im nächsten Kapitel detailliert definieren, . Wir erhalten folgendes Resultat:

Proposition 2.3.1. *Es seien $\boldsymbol{\pi}^\varepsilon = (\pi^\varepsilon_x)_{x\in\mathcal{S}}$ und $\boldsymbol{\pi}^{\mathbf{Q}} = (\pi^{\mathbf{Q}}_i)_{i=1}^n$ die stationären Verteilungen von Z^ε bzw. Y. Dann gilt*

(i) $\lim_{\varepsilon\to 0} \pi^\varepsilon_x = 0$ falls $x \neq \mathfrak{m}_i$ für alle $1 \leq i \leq n$.

(ii) $\lim_{\varepsilon\to 0} \pi^\varepsilon_{\mathfrak{m}_i} = \pi^{\mathbf{Q}}_i$ für $1 \leq i \leq n$.

Im dritten Kapitel widmen wir uns dem Studium der *Metastabilität*. Dieser Begriff bezieht sich dabei auf die Existenz von Zeitskalen, auf der man unterschiedliche statistische Gleichgewichte für die Lösung von Gleichung (0.0.1) beobachten kann, denn in Abhängig

von der Art der Störung sowie den Startbedingungen kann es in der Tat passieren, dass bestimmte Gebiete erst nach einer verhältnismäßig langen Zeit erreicht werden oder gar gänzlich unberührt bleiben. Grob gesprochen bedeutet dies, dass der Prozess X^ε bei zu kurzem Zeithorizont der deterministischen Trajektorie folgt und zu dem lokalen Minimum des Potentialtopfes konvergiert, in dem es auch gestartet ist, während die Wahl von ausreichend langen Zeitskalen dazu führen kann, dass sich X^ε annähernd wie ein Markov-Sprungprozess mit Zustandsraum $\{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}$ verhält. Kipnis und Newman betrachteten in [KiNe] beispielsweise den einfachen Fall, dass U ein Zweitopfpotential, d.h. es gibt 2 Minima $\mathfrak{M}$, $\mathfrak{m}$ und ein Maximum $\mathfrak{s}$ mit $U(\mathfrak{M}) < U(\mathfrak{m}) < U(\mathfrak{s})$, und η_t eine Brownsche Bewegung ist, und konnten zeigen, dass X^ε auf einer mit einem Faktor ρ^ε skalierten Zeitachse gegen einen Markov-Prozess mit zwei Zuständen konvergiert, bei welchem das globale Minimum $\mathfrak{M}$ absorbierend ist. Dabei wächst dieser Skalierungsfaktor exponentiell schnell bezüglich ε und die logarithmische Rate ist gegeben durch

$$\lim_{\varepsilon \to 0} \varepsilon^2 \ln \rho^\varepsilon = 2(U(\mathfrak{s}) - U(\mathfrak{m})).$$

Wie man sieht, hängt diese Rate von der Höhe des flachen Potentialtopfes ab. Darüber hinaus ist ρ^ε durch die mittlere Austrittszeit aus diesem Topf gegeben. Das mehrdimensionale Analogon zu dieser Situation, d.h. wo $U' = \nabla F$ und $F\colon \mathbb{R}^d \to \mathbb{R}$ ein glattes Vektorfeld mit zwei lokalen Minima und einem Sattelpunkt ist, wurde in [GOV] betrachtet.

Häufig wird der Gaußsche Fall mit Hilfe der Theorie über große Abweichungen behandelt. Diese wird in dem Buch [FrWe] von Freidlin und Wentzell vorgestellt und impliziert, dass, wenn Ω_i und Ω_j zwei verschiedene Potentialtöpfe sind, es Zahlen $c_{i,j}$ derart gibt, dass sich die erwartete Übergangszeit von Ω_i nach Ω_j asymptotisch wie $\mathrm{e}^{c_{i,j}/\varepsilon^2}$ verhält. Im nicht-entarteten Fall, d.h. alle Potentialtöpfe besitzen eine unterschiedliche Tiefe, sind diese Zahlen für jede Paarung i, j verschieden und folglich können diese Übergangszeiten, obgleich alle exponentiell groß, getrennt werden. Dies führt für gegebenem "Starttopf" Ω_i zu einer Hierarchie von mittleren Übergängen, der sogenannten *Hierarchie von Zyklen.* Wir verweisen für weitere Details auf das Buch [FrWe] sowie einen neueren Artikel von Cameron [Ca]. Zusätzlich möchten wir erwähnen, dass ein potentialtheoretische Ansatz zur Metastabilität im Gaußschen Fall in dem Paper [BEGK1] diskutiert wird.

Verwendet man wie in Gleichung (0.0.1) als Rauschterm η_t allerdings symmetrische α-stabile Lévy-Prozesse, so kann man ein anderes Verhalten beobachten. Beispielsweise verhindert die Möglichkeit großer instantaner Sprünge, dass X^ε in einem Topf "gefangen" ist, also dass es ein Minimum gibt, in dessen Umgebung X^ε absorbiert wird. Das metastabile Verhalten für diesen Fall wurde ausführlich von Imkeller und Pavlyukevich in [ImPa2] untersucht. Sie konnten beweisen, dass X^ε auf einer durch den Faktor $\frac{\alpha}{2\varepsilon^\alpha}$ skalierten Zeitachse im Sinne endlichdimensionaler Verteilungen gegen einen Markov-Prozess Y mit Zustandsraum $\{\mathfrak{m}_1, \dots, \mathfrak{m}_n\}$ konvergiert, welcher durch die Matrix $\mathbf{Q} = (q_{i,j})_{i,j=1}^n$ mit

$$q_{i,j} = \begin{cases} \frac{1}{2}\left| \frac{1}{|\mathfrak{s}_{j-1}-\mathfrak{m}_i|^\alpha} - \frac{1}{|\mathfrak{s}_j-\mathfrak{m}_i|^\alpha} \right|, & i \neq j, \\ -\frac{1}{2}\left(\frac{1}{|\mathfrak{s}_i-\mathfrak{m}_i|^\alpha} + \frac{1}{|\mathfrak{s}_i-\mathfrak{m}_i|^\alpha} \right), & i = j, \end{cases}$$

erzeugt wird. Wie man sehen kann, spielt im Gegensatz zum Gaußschen Fall nicht die Höhe der Ω_i sondern deren Ausdehnung eine wichtige Rolle.

Entscheidend in der Herleitung dieses Resultates ist die Untersuchung der Austrittszeiten aus den Potentialtöpfen, von denen die Autoren zeigen, dass sie von Ordnung $O(\varepsilon^{-\alpha})$ sind und nach geeigneter Skalierung im Grenzwert $\varepsilon \to 0$ gegen eine exponentialverteilte Zufallsgröße konvergieren. Eine wichtige Vorarbeit dafür wurde bereits in der Arbeit [ImPa1] geleistet, wobei dort noch nicht das Verhalten in der Nähe eines Sattelpunktes betrachtet wurde. Wie sich herausstellt, führt der Mangel an treibenden (deterministische) Kräften in den Umgebungen der lokalen Maxima dazu, dass X^ε dazu tendiert, sich verhältnismäßig lange dort aufzuhalten.

Da die in Kapitel 2 definierte Markov-Kette Z^ε eine Approximation von X^ε darstellt, kann man ein vergleichbares metastabiles Verhalten erwarten. Im Zentrum des Kapitels steht daher der Beweis des folgenden Theorems:

Theorem 3.0.1. *Es existiert eine Konstante $c(\alpha, h) > 0$ derart, dass für jedes $t > 0$ und alle $1 \leq i, j \leq n$ mit $i \neq j$ gilt*

$$\lim_{\varepsilon \to 0} \mathbb{P}_x \left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t \rceil} = \mathfrak{m}_j \right) = \mathbb{P}_{\mathfrak{m}_i}(Y_t = \mathfrak{m}_j)$$

gleichmäßig in $x \in \mathcal{S}$.

Auch wir untersuchen dafür zunächst die Austrittszeiten aus einem der Potentialtöpfe und

erhalten ein zu [ImPa2] analoges Resultat, d.h diese Zeiten als auch deren Mittelwerte sind von Ordnung $O(\varepsilon^{-\alpha})$. Durch die Tatsache, dass die lokalen Maxima nach Definition des Zustandsraumes keine Zustände von Z^ε darstellen, sind wir aber im Gegensatz zu den Autoren in [ImPa2] nicht gezwungen, der Dynamik in der Nähe der Sattelpunkte besondere Aufmerksamkeit zu schenken. Zusätzlich studieren wir die erwarteten Übergangszeiten zwischen zwei fest gewählten Potentialtöpfen und zeigen, dass auch diese polynomiell in $\varepsilon^{-\alpha}$ wachsen.

Zusätzlich sei angemerkt, dass metastabiles Verhalten für zeitdiskrete Markov-Ketten in einem abstrakteren Kontext unter anderem auch in den Arbeiten [BEGK2] und [Eck] untersucht wird. Der nennenswerte Unterschied zu der hier vorliegenden Arbeit ist zum einen die zusätzliche Forderung der Reversibilität in [BEGK2], welche bei uns nicht gegeben ist, und zum anderen die Voraussetzung, dass man die Potentialtöpfe hinsichtlich ihrer "Anziehungskraft", ausgedrückt durch erwartete Übergangszeiten, unterscheiden kann. Dies ist beispielsweise im Falle eines Gaußschen Rauschterms möglich. In unserem Szenario ist das allerdings nicht gegeben, da die mittleren Übergangszeiten alle dieselbe polynomielle Ordnung $O(\varepsilon^{-\alpha})$ besitzen.

Das vierte Kapitel widmet sich dann der Spektralanalyse für den infinitesimalen Erzeuger (bzw. dessen diskretem Analogon) der Markov-Kette Z^ε. Betrachten wir zunächst wieder Gleichung (0.0.1) mit einer Gaußschen Störung. Der infinitesimale Erzeuger hat dann für hinreichend glatte Testfunktionen die Gestalt

$$(\boldsymbol{\mathcal{D}}^\varepsilon f)(x) = -U'(x)\frac{d}{dx}f(x) + \frac{\varepsilon^2}{2}\frac{d^2}{dx^2}f(x)$$

und ist selbstadjungiert auf dem gewichteten L_2-Raum $L_2(\mathrm{e}^{-\frac{2U(x)}{\varepsilon^2}}\,dx)$. Das Spektralverhalten dieses Operators wird seit Jahrzehnten intensiv studiert, siehe unter anderem die Arbeiten [BGK], [BuMa1], [BuMa2], [FrWe] und [KoMa], und kann wie folgt kurz zusammengefasst werden: Es gibt n ausgezeichnete Eigenwerte $\lambda_1^\varepsilon, \ldots, \lambda_n^\varepsilon$. Dabei ist $\lambda_1^\varepsilon = 0$ und die restlichen Werte $\lambda_i^\varepsilon,\ 2 \leq i \leq n$, sind von exponentiell kleiner Ordnung, $\lambda_i^\varepsilon = O\left(\mathrm{e}^{-c_i\varepsilon^2}\right)$, und durch eine Spektrallücke vom Rest des Spektrums getrennt. Die zahlen c_i hängen wieder von den Höhen der Potentialtöpfe ab. Ist U also ein nicht-entartetes Potential,

so kann man diese Eigenwerte anhand dieser logarithmischen Raten "trennen" . Eine für diesen Fall analoge diskrete Situation wurde in [BEGK2] untersucht. Außerdem kann man diese Eigenwerte mit den inversen mittleren Austrittszeiten aus den Potentialtöpfen identifizieren. In der Tat, Kolokoltsov und Makarov studierten in [KoMa] (siehe auch [Ko2], Kapitel 8, Sektion 2) zwei unterschiedliche Situationen. Sei δ^j die Indikatorfunktion des Topfes Ω_j. Sie betrachteten im Fall A ein kompaktes Intervall M, welches alle lokalen Minima von U enthält, und das inhomogene Differentialgleichungssystem

$$\boldsymbol{\mathcal{D}}^\varepsilon \tau^{j,\varepsilon} = \delta^j, \quad \tau^{j,\varepsilon}|_{M^c} = 0, \quad 1 \leq j \leq n.$$

Im Fall B betrachteten sie das Problem auf der gesamten reellen Achse. Dann hat das Differentialgleichungssystem die Form

$$\boldsymbol{\mathcal{D}}^\varepsilon \tau^{j,\varepsilon} = \delta^j - \pi^\varepsilon(\Omega_j)\mathbf{1}, \quad 1 \leq j \leq n, \tag{0.0.2}$$

wobei π^ε die stationäre Verteilung von X^ε bezeichnet und $\mathbf{1}(x) = 1$ für alle $x \in \mathbb{R}$. Es ist bekannt, dass im Fall A die Lösung $\tau^{j,\varepsilon}(x)$ der Zeit entspricht, die der Prozess in Ω_j verbringt wenn er in x gestartet ist. Sei nun $\mathbf{G}^\varepsilon = (g^\varepsilon_{i,j})^n_{i,j=1}$ die Matrix, die die Mittelungen (bzgl. π^ε) der Lösungen $\tau^{j,\varepsilon}$ über Ω_i enthält, d.h.

$$g^\varepsilon_{i,j} := \frac{1}{\pi^\varepsilon(\Omega_i)} \int_{\Omega_i} \tau^{j,\varepsilon}(x)\pi^\varepsilon(dx).$$

Falls λ^ε_i, $1 \leq i \leq n$, die oben genannten kleinen Eigenwerte von $\mathcal{D}^\varepsilon$ und μ^ε_i die Eigenwerte von $\mathbf{G}^\varepsilon$ sind, dann besagt Theorem 1.2 in [KoMa], dass die Beziehung $\mu^\varepsilon_i = (\lambda^\varepsilon_i)^{-1}(1 + O(\varepsilon^\infty))$ gilt (im Fall B hat man zusätzlich $\mu^\varepsilon_1 = \lambda^\varepsilon_1 = 0$).

Darüber hinaus weiß man, dass die zugehörigen Eigenfunktion bzw. Eigenvektoren über den Anziehungsgebieten näherungsweise konstant sind.

Ist die Störung in Gleichung (0.0.1) ein symmetrischer α-stabiler Lévy-Prozess, dann ist der Erzeuger der Lösung X^ε ein nicht-symmetrischer Integro-Differentialoperator und hat die Gestalt

$$(\mathcal{A}^\varepsilon f)(x) = -U'(x)\frac{d}{dx}f(x) + c(\alpha)\varepsilon^\alpha \int_{\mathbb{R}\setminus\{0\}} \frac{f(x+y) - f(x) - yf'(x)\mathbf{1}_{\{|y|\leq 1\}}(y)}{|y|^{\alpha+1}} dy.$$

Die Vermutung ist, dass es auch hier eine Menge von n Eigenwerten $\lambda^\varepsilon_1 = 0, \ldots, \lambda^\varepsilon_n$ gibt, die mittels einer Spektrallücke vom Restspektrum abgespalten werden können. Die enge

Verbindung zu mittleren Austrittszeiten und die Tatsache, dass diese von Ordnung $O(\varepsilon^{-\alpha})$ sind (siehe [ImPa1], [ImPa2]), lassen den Schluss zu, dass diese Eigenwerte von polynomiell kleiner Ordnung $O(\varepsilon^{\alpha})$ sind. Der Spektralanalyse für diesen Operator erweist sich allerdings als sehr kompliziert, weswegen wir uns in dieser Arbeit der Einfachheit halber auf die diskrete Situation mit der Markov-Kette Z^{ε} zurückziehen, welche als Approximation von X^{ε} angesehen werden kann. Deshalb bleibt die dargelegte Heuristik unverändert, was allerdings zu Problemen führt. Genauer gesagt: Die unterschiedlichen logarithmischen Raten im Gaußschen Fall erlauben eine Unterscheidung der Eigenwerte, obwohl diese allesamt exponentiell klein sind. In der analogen diskreten Situation, welche in [BEGK2] und [Eck] studiert wird, ist dies durch das Prinzip der Nicht-Entartung gegeben, welche mit Hilfe von mittleren Übergangszeiten definiert wird. Letztere sind wiederum eng mit den Eigenwerten verbunden. Durch unsere Vermutung, dass die Eigenwerte von derselben polynomiell kleinen Ordnung sind, werden wir gezwungen, andere Methoden zu benutzen. Der Ansatz besteht darin, die Kenntnisse aus Kapitel 3 über die "richtige" Zeitskala zu verwenden. Für endliche, zeitstetige Markov-Ketten ist die Beziehung zwischen der Familie $(\mathbf{P}_t)_{t\geq 0}$ von Übergangsmatrizen und dem infinitesimalen Erzeuger $\mathbf{Q}$ gegeben durch

$$\mathbf{P}_t = \text{“ } \mathrm{e}^{t\mathbf{Q}} \text{ ”} = \sum_{k=0}^{\infty} \frac{t^k \mathbf{Q}^k}{k!}.$$

Im zeitdiskreten Fall gilt die Gleichung $\mathbf{Q} = \mathbf{P} - \mathbf{I}$. Nach Theorem 3.0.1 verhält sich Z^{ε} im Grenzwert $\varepsilon \to 0$ auf der durch den Streckungsfaktor $\frac{c(\alpha,h)}{\varepsilon^{\alpha}}$ skalierten Zeitachse statistisch annähernd wie der Markov-Sprungprozess Y mit Erzeuger $\mathbf{Q}$. Die Vermutung ist also, dass ein Teil des Spektrums der Matrix $\mathbf{Q}^{\varepsilon} = \frac{c(\alpha,h)}{\varepsilon^{\alpha}}(\mathbf{P}^{\varepsilon} - \mathbf{I})$ gegen das Spektrum $\sigma(\mathbf{Q}) = \left\{0, \lambda_2^{\mathbf{Q}}, \ldots, \lambda_n^{\mathbf{Q}}\right\}$ von $\mathbf{Q}$ konvergiert. Es gilt folgendes Theorem:

Theorem 4.2.2. *Das Spektrum $\sigma(\mathbf{Q}^{\varepsilon})$ kann in zwei disjunkte Teilmengen $\sigma_1(\mathbf{Q}^{\varepsilon})$ und $\sigma_2(\mathbf{Q}^{\varepsilon})$ geteilt werden. Für diese gelten die folgenden Aussagen:*

(i) $\sigma_1(\mathbf{Q}^{\varepsilon})$ enthält genau n Eigenwerte $\lambda_1^{\varepsilon}, \ldots, \lambda_n^{\varepsilon}$. Dabei gilt $\lambda_1^{\varepsilon} = 0$ und

$$\lim_{\varepsilon \to 0} \lambda_i^{\varepsilon} = \lambda_i^{\mathbf{Q}}, \quad 2 \leq i \leq n.$$

(ii) Für die restlichen Eigenwerte $\lambda_{n+1}^{\varepsilon}, \dots, \lambda_N^{\varepsilon}$ *gilt*

$$\lim_{\varepsilon \to 0} |\lambda_i^{\varepsilon}| = \infty, \quad n+1 \leq i \leq N.$$

Insbesondere folgt daraus, dass es eine Spektrallücke gibt, welche $\sigma_1(\mathbf{Q}^{\varepsilon})$ *von* $\sigma_2(\mathbf{Q}^{\varepsilon})$ *trennt.*

Der entscheidende Schritt im Beweis von Theorem 4.2.2 ist die Analyse des Konvergenzverhaltens des charakteristischen Polynoms von $\mathbf{Q}^{\varepsilon}$ bezüglich dessen wir folgenden Satz beweisen.

Proposition 4.2.3. *Es bezeichnen* P^{ε} *und* $P^{\mathbf{Q}}$ *die charakteristischen Polynome von* $\mathbf{Q}^{\varepsilon}$ *beziehungsweise* $\mathbf{Q}$*. Dann gibt es ein* $K \geq 1$ *sowie Konstanten* $c(\alpha, h) > 0$ *und* $C > 0$*, welche ebenfalls von den Diskretisierungsparametern und* α *abhängen kann, so dass für jedes* $0 < \varepsilon < \varepsilon_0$ *und jedes* $\lambda \in \mathbb{C}$

$$\left| \left(-\frac{\varepsilon^{\alpha}}{c(\alpha, h)} \right)^{N-n} P^{\varepsilon}(\lambda) - P^{\mathbf{Q}}(\lambda) \right| \leq C |\lambda|^K \varepsilon^{\alpha}.$$

In Abschnitt 4.3 widmen wir uns dann dem Studium der Eigenvektoren und bestätigen die Vermutung, dass diese über den Anziehungsgebieten annähernd konstant sind. Sei $\boldsymbol{\delta}^j = (\delta_x^j)_{x \in \mathcal{S}}$ der Indikatorvektor von $\mathcal{S} \cap [\mathfrak{s}_{j-1}, \mathfrak{s}_j]$, d.h. $\delta_x^j = 1$ falls $x \in \mathcal{S} \cap [\mathfrak{s}_{j-1}, \mathfrak{s}_j]$ und 0 sonst. Zusätzlich nehmen wir an, dass, für alle $1 \leq i \leq n$, die Eigenräume zu den Eigenwerten $\lambda_i^{\mathbf{Q}}$ eindimensional sind. Dann gilt

Theorem 4.3.1. *Es seien* $\lambda_1^{\varepsilon}, \dots, \lambda_n^{\varepsilon}$ *die n Eigenwerte aus Theorem 4.2.2 (i) und* $\boldsymbol{\psi}^{\varepsilon,i} = (\psi_x^{\varepsilon,i})_{x \in \mathcal{S}}$ *die zugehörigen Rechtseigenvektoren. Diese seien derart normiert, dass* $\max_{x \in \mathcal{S}} |\psi_x^{\varepsilon,i}| = 1$*. Ferner sei der Rechtseigenvektor* $\boldsymbol{\psi}^{\mathbf{Q},i} = (\psi_j^{\mathbf{Q},i})_{j=1}^n$ *zu* $\lambda_i^{\mathbf{Q}}$ *derart normiert, dass* $\max_{1 \leq j \leq n} \left| \psi_j^{\mathbf{Q},i} \right| = 1$*. Dann gilt* $\boldsymbol{\psi}^{\varepsilon,1} = (1, \dots, 1)$ *und*

$$\lim_{\varepsilon \to 0} \left| \psi_x^{\varepsilon,i} - \sum_{j=1}^{n} \psi_j^{\mathbf{Q},i} \delta_x^j \right| = 0, \quad x \in \mathcal{S}, \quad 2 \leq i \leq n.$$

Die Herleitung dieses Resultates folgt den Ideen aus den Arbeiten [BEGK2] und [Eck].

Aufbauend auf den Ergebnissen aus den Kapiteln 3 und 4 untersuchen wir abschließend die

Beziehung der Eigenwerte aus Theorem 4.2.2 (i) und mittleren Austrittszeiten. Motiviert wird dies durch die Arbeiten [KoMa] bzw. [Ko2], Kapitel 8, in welcher, wie bereits erwähnt, der Gaußsche Fall studiert wird. Das diskrete Analogon zum Differentialgleichungssystem (0.0.2) sind die linearen Gleichungssysteme

$$(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\boldsymbol{\tau}^{\varepsilon,j} = \boldsymbol{\pi}^\varepsilon(\mathcal{S}_j)\mathbf{1} - \boldsymbol{\delta}^j, \quad 1 \le j \le n.$$

Mit Hilfe der Lösungen $\boldsymbol{\tau}^{\varepsilon,j}$ dieser Systeme definieren wir eine Matrix $\mathbf{G}^\varepsilon = (g^\varepsilon_{i,j})^n_{i,j=1}$ mittels

$$g^\varepsilon_{i,j} := \frac{\sum_{x\in\mathcal{S}_i} \tau^{\varepsilon,j}_x \pi^\varepsilon_x}{\sum_{x\in\mathcal{S}_i} \pi^\varepsilon_x}.$$

Dann erhält man unter der Annahme einer zusätzlichen, rein technischen Bedingung (B) das folgende Theorem.[2]

Theorem 5.0.1 *Es seien $\lambda^\varepsilon_1, \dots, \lambda^\varepsilon_n$ die Eigenwerte von $\mathbf{Q}^\varepsilon$ aus Theorem 4.2.2(i) und $\mu^\varepsilon_1, \dots, \mu^\varepsilon_n$ seien die Eigenwerte von $\overline{\mathbf{G}}^\varepsilon := \frac{\varepsilon^\alpha}{c(\alpha,h)}\mathbf{G}^\varepsilon$. Unter der Annahme, dass Bedingung (B) erfüllt ist, folgt*

$$\mu^\varepsilon_1 = \lambda^\varepsilon_1 = 0 \quad \textit{und} \quad \lim_{\varepsilon\to 0}\left|(\lambda^\varepsilon_i)^{-1} - \mu^\varepsilon_i\right| = 0, \quad 2 \le i \le n.$$

Im Gegensatz zu der Arbeit [KoMa] können wir dieses Theorem nicht "direkt" beweisen. Glücklicherweise stellt sich heraus, dass man Lösungen $\boldsymbol{\tau}^j$ obiger Gleichungssysteme mit Hilfe der sogenannten *Fundamentalmatrix* $\mathbf{F}^\varepsilon = (f^\varepsilon_{x,y})_{x,y\in\mathcal{S}}$ bzw. *Abweichungsmatrix* $\mathbf{D}^\varepsilon = (d^\varepsilon_{x,y})_{x,y\in\mathcal{S}}$ konstruieren kann. Eingeführt wurde die Fundamentalmatrix von Kemeny und Snell in [KeSn1] bzw. [KeSn2]. Für zeitdiskrete Markov-Ketten ist sie definiert durch

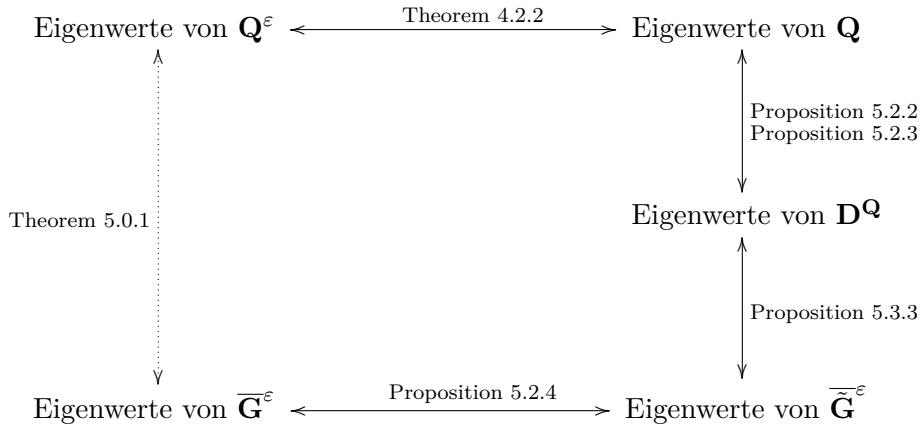
$$\mathbf{F}^\varepsilon := (\mathbf{I} - (\mathbf{P}^\varepsilon - \mathbf{\Pi}^\varepsilon))^{-1}$$

mit $\mathbf{\Pi}^\varepsilon := (\pi^\varepsilon_y)_{x,y\in\mathcal{S}}$. Die Abweichungsmatrix ist dann gegeben durch $\mathbf{D}^\varepsilon := \mathbf{F}^\varepsilon - \mathbf{\Pi}^\varepsilon$. Über diese konstruieren wir eine weitere Matrix $\tilde{\mathbf{G}}^\varepsilon = (\tilde{g}^\varepsilon_{i,j})^n_{i,j=1}$ mittels

$$\tilde{g}^\varepsilon_{i,j} := \frac{\sum_{x\in\mathcal{S}_i} \pi^\varepsilon_x \left(\sum_{y\in\mathcal{S}_j} d^\varepsilon_{x,y}\right)}{\pi^\varepsilon(\mathcal{S}_j)}.$$

[2] Die konkrete Formulierung dieser Bedingung erfolgt in Kapitel 5.

Wir zeigen in Proposition 5.2.4, dass diese Matrix (im wesentlichen) die gleichen Eigenwerte besitzt wie $\mathbf{G}^{\varepsilon}$. Bezeichnet man mit $\mathbf{D}^{\mathbf{Q}}$ die Abweichungsmatrix der Kette Y und definiert $\overline{\widetilde{\mathbf{G}}}^{\varepsilon} := \frac{\varepsilon^{\alpha}}{c(\alpha,h)}\tilde{\mathbf{G}}^{\varepsilon}$, so kann man die Zusammenhänge zwischen den Eigenwerten dieser Matrizen, und damit die Beweisstruktur von Theorem 5.0.1, mit dem folgenden Schema veranschaulichen:



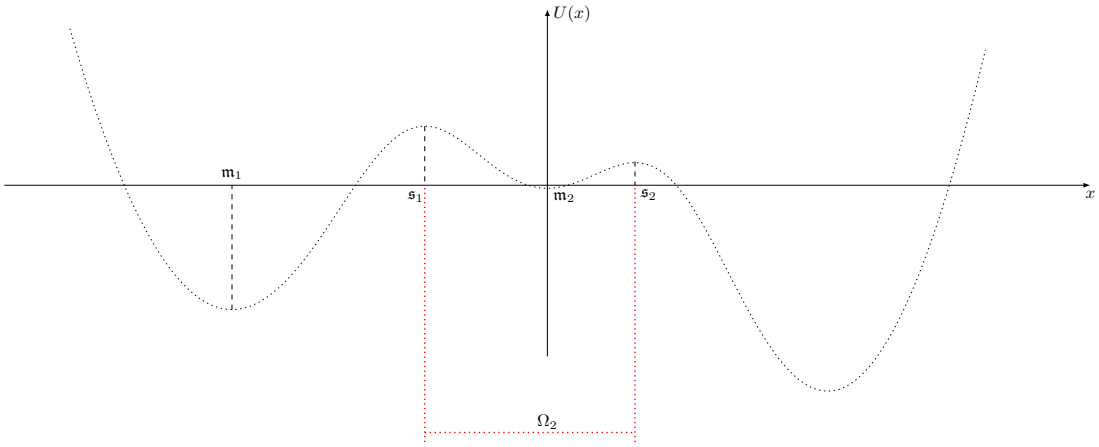
Dabei wird im Beweis von Proposition 5.3.2 die enge Verbindung der Abweichungsmatrix mit der Matrix $\mathbf{M} = (m_{x,y})_{x,y\in\mathcal{S}}$ von mittleren Übergangszeiten ausgenutzt, auf welche wir die Ergebnisse aus Kapitel 3 anwenden können.

Introduction and Summary of the Results

As the point of origin we consider the one-dimensional deterministic dynamical system

$$\dot{X}_t = -U'(X_t), \quad t \geq 0,$$

with the initial condition $X_0 = x$. Here $U\colon \mathbb{R} \to \mathbb{R}$ is a multi-well potential with local minima $\mathfrak{m}_1, \ldots, \mathfrak{m}_n,\ n \geq 1$, as well as $n-1$ local maxima $\mathfrak{s}_1, \ldots, \mathfrak{s}_{n-1}$. The potential wells Ω_i are the intervals between two local maxima, i.e. $\Omega_i = (\mathfrak{s}_{i-1}, \mathfrak{s}_i)$.



Physically it is clear that the trajectory $(X_t)_{t\geq 0}$ of a particle in the force field generated by U either converges to a local minimum or stays in an extremum of U, depending on the initial value. A transition between different wells is not possible.
This situation changes if we add a small (stochastic) noise in the equation above. One then derives the stochastic differential equation

$$X_t^\varepsilon = x - \int_0^t U'(X_s^\varepsilon)ds + \varepsilon\eta_t \tag{0.0.1}$$

with a stochastic process $\eta = (\eta_t)_{t\geq 0}$ as the noise term of which the intensity is controlled by the parameter $\varepsilon > 0$. Unquestionably, the case that is considered most is $\eta = B$ where B is a Brownian motion. In literature this case is normally referred to as *diffusion*.[3]
Diffusion equations are intensively studied for decades and are used for modeling in physics, biology, financial mathematics and so forth. Standard literature on the theory of stochastic differential equations are, for example, the books [Ok] and [KaSh] while the main reference for the small noise limit, i.e. equation (0.0.1) in the limit $\varepsilon \to 0$, is the book [FrWe] by Freidlin and Wentzell. A disadvantage in using a Brownian motion, which is also known as Gaussian *white noise* in the context of stochastic differential equations, as the noise in models is that the occurrence of *extreme events* such as crashes in stock markets or natural catastrophes cannot be modeled accurately since the tails of the normal distribution decrease exponentially fast. Therefore, in recent years scientists started to consider equations of the form (0.0.1) with noise terms η which allow the occurrence of instantaneous jumps. An important class of such stochastic processes is given by the set of symmetric α-stable Lévy processes where $0 < \alpha < 2$ is the so-called index of stability. These are often referred to as *Lévy noise* or, in the context of modeling of physical problems, as *Lévy Flights*.
A concrete example for a geophysical phenomenon that is modeled by equation (0.0.1) with a symmetric α-stable Lévy noise and a multi-well potential U is presented in papers of Ditlevsen, see [Di1] and [Di2]. The author studied the evolution of the temperature in the Northern hemisphere over the past 100 000 years by analyzing the data received from ice-core measurements. Reconstructions based on these measurement indicate that there were about 25 abrupt climatic changes which can be interpreted as the aforementioned "catastrophic" events. Indeed, an easy model to describe these observations would be the choice of a two-well potential where the local minima correspond to the states "warm" and "cold".

Now we also consider equation (0.0.1) driven by a symmetric α-stable Lévy noise. This

[3] Let us remark that, in general, the term *diffusion* includes more general drifts than just multi-well potentials. Also, the external control of the intensity by a parameter ε is often omitted.

thesis is devoted to the study of a Markov chain that is induced by a discretization scheme applied to this equation.

The first chapter begins with an overview of the notations and definitions that we are going to use. We start off with Laplace transforms of random variables although a slightly different definition compared to the standard literature is used. Afterwards we will present the definitions and results from the theory of Markov processes that are needed in this thesis and the last part is devoted to a more detailed introduction of symmetric α-stable Lévy processes.

The second chapter is devoted to the introduction of two discrete-time Markov chains Z^0 and Z^ε on a common finite state space $\mathcal{S}$. These chains are supposed to function as discretizations of the solutions of equation (0.0.1). With that in mind we first choose an Euler scheme with step size $h > 0$ to discretize the time axis. Using the self-similarity of Lévy processes this procedure yields the recursion equation

$$\tilde{X}_0 = x, \quad \tilde{X}^\varepsilon_{kh} = \tilde{X}^\varepsilon_{(k-1)h} - hU'(\tilde{X}^\varepsilon_{(k-1)h}) + h^{\frac{1}{\alpha}}\varepsilon\xi_k, \quad k \geq 1,$$

where $(\xi_k)_{k\geq 1}$ is a sequence of independent and identically distributed random variables such that $\xi_1 \stackrel{d}{=} L_1$. To define the state space $\mathcal{S}$ we restrict our considerations on a compact interval $[-R, R]$ with an $R > 0$ large enough such that it contains all local minima of U. Afterwards we divide this interval in N disjoint sub-intervals which need not to be equidistant but also cannot exceed a length of $\delta > 0$. Then we choose exactly one point out of each of these intervals under the additional rule that the local minima must be chosen and the local maxima must not. Altogether we then have N points that form the state space $\mathcal{S}$. For any state $y \in \mathcal{S}$ denote by I_y the subinterval that contains y. Then the transition probabilities $p^0_{x,y},\ x, y \in \mathcal{S}$, for Z^0 and $p^\varepsilon_{x,y}$ for Z^ε are defined by

$$p^0_{x,y} := \mathbb{P}(x - hU'(x) \in I_y) \quad \text{and} \quad p^\varepsilon_{x,y} := \mathbb{P}(x - hU'(x) + h^{\frac{1}{\alpha}}\varepsilon\xi_1 \in I_y), \text{ respectively.}$$

Using a known series expansion for the tails $\mathbb{P}(\xi_1 > a)$, $a \to \infty$, we are able to prove the following statements about the transition probabilities.

Proposition 2.2.1. *Let $\mathbf{P}^0 = (p^0_{x,y})_{x,y\in\mathcal{S}}$ and $\mathbf{P}^\varepsilon = (p^\varepsilon_{x,y})_{x,y\in\mathcal{S}}$ be the transition matrices for Z^0 and Z^ε, respectively. Then:*

(i) For every $x \in \mathcal{S}$ there is precisely one state $y^* = y^*(x)$ such that $p^0_{x,y^*} = 1$.

(ii) Let $x \in \mathcal{S}$ and let y^* be the uniquely determined state from (i). Then $1 - p^\varepsilon_{x,y^*} = O(\varepsilon^\alpha)$ and $p^\varepsilon_{x,y} = O(\varepsilon^\alpha)$ for all remaining states $y \neq y^*$.

In the last part of this section we investigate the asymptotic behavior of the stationary distribution of Z^ε. For this we have to forestall a quantity, namely a continuous-time Markov chain Y with state space $\{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}$, which will be defined in more detail in the next chapter.

Proposition 2.3.1. *Let $\boldsymbol{\pi}^\varepsilon = (\pi^\varepsilon_x)_{x\in\mathcal{S}}$ and $\boldsymbol{\pi}^{\mathbf{Q}} = (\pi^{\mathbf{Q}}_i)_{i=1}^n$ be the stationary distributions of Z^ε and Y, respectively. Then:*

(i) $\lim_{\varepsilon\to 0} \pi^\varepsilon_x = 0$ if $x \neq \mathfrak{m}_i$ for all $1 \leq i \leq n$.

(ii) $\lim_{\varepsilon\to 0} \pi^\varepsilon_{\mathfrak{m}_i} = \pi^{\mathbf{Q}}_i$ for $1 \leq i \leq n$.

The third chapter is devoted to the study of a phenomenon called *metastability*. This term refers to the existence of a time scale on which it is possible to observe different statistical equilibriums for the solution of equation (0.0.1). Indeed, depending on the nature of the noise term and the initial state it is possible that certain potential wells can only be reached after a relatively long time or stay unvisited entirely. Roughly spoken that means that if the time horizon is too short, then the process X^ε will follow the deterministic trajectory and tend to the local minimum of the potential well where it started. But if one chooses a sufficiently long time scale it is possible that the process will behave like a Markov jump process on the state space $\{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}$. For example Kipnis and Newman considered in [KiNe] the simple case where U is a two-well-potential, i.e. there are two minima $\mathfrak{m}$ and $\mathfrak{M}$ and a maximum $\mathfrak{s}$ such that $U(\mathfrak{M}) < U(\mathfrak{m}) < U(\mathfrak{s})$, and where η is a Brownian motion. They showed that on a time scale stretched by a factor ρ^ε the solution X^ε converges to a Markov process with two states of which the global minimum $\mathfrak{M}$ is absorbing. Here the scaling factor increases exponentially fast with respect to ε and the logarithmic rate is given by

$$\lim_{\varepsilon\to 0} \varepsilon^2 \ln \rho^\varepsilon = 2(U(\mathfrak{s}) - U(\mathfrak{m})).$$

As one can see this rate depends on the depth of the shallow potential well. Moreover, one can prove that ρ^ε is determined by the mean exit time from this well. The multi-dimensional analog, i.e. where $U' = \nabla F$ and $F\colon \mathbb{R}^d \to \mathbb{R}$ is a smooth vector field with three critical points (two minima and one saddle point), is considered in [GOV].

In general, the Gaussian case is often treated with the help of the large deviations theory given by Freidlin and Wentzell [FrWe] which implies that if Ω_i and Ω_j are two different wells it can be shown that there is a number $c_{i,j}$ such that the expected transition time from Ω_i to Ω_j behaves asymptotically like $\mathrm{e}^{c_{i,j}/\varepsilon^2}$. In the generic case, i.e. the potential wells of U possess different depths, these numbers $c_{i,j}$ are different for each pairing i, j and hence these transition times, although all of them are exponentially large, can be separated. That leads in a natural way to a hierarchy of expected transitions given a "starting well" Ω_i, the so-called *hierarchy of cycles*. For more details we refer to [FrWe] as well as a recent work by Cameron [Ca]. Let us also mention that a potential theoretic approach for metastability in the Gaussian case is presented in [BEGK1].

If we use an α-stable Lévy noise in equation (0.0.1), then a different behavior is observed. The existence of large instantaneous jumps, for example, forbids X^ε to be trapped in a well, i.e. that there exists a local minimum and a neighborhood around this minimum in which X^ε will be absorbed eventually. The metastable behavior for this case was intensively studied in [ImPa2]. The authors proved that on a time scale stretched by the factor $\frac{\alpha}{2\varepsilon^\alpha}$ the solution converges in the sense of finite dimensional distributions to a Markov process Y with state space $\{\mathfrak{m}_1, \dots, \mathfrak{m}_n\}$ generated by the matrix $\mathbf{Q} = (q_{i,j})_{i,j=1}^n$ where the entries are given by

$$q_{i,j} = \begin{cases} \frac{1}{2}\left|\frac{1}{|\mathfrak{s}_{j-1}-\mathfrak{m}_i|^\alpha} - \frac{1}{|\mathfrak{s}_j-\mathfrak{m}_i|^\alpha}\right|, & i \neq j, \\ -\frac{1}{2}\left(\frac{1}{|\mathfrak{s}_i-\mathfrak{m}_i|^\alpha} + \frac{1}{|\mathfrak{s}_i-\mathfrak{m}_i|^\alpha}\right), & i = j. \end{cases}$$

Here one can observe that, in contrast to the Gaussian case, the width of the potential wells Ω_i is crucial and not their depths.

The derivation of this result is mainly based on the extensive investigation of the mean exit times of the potential wells. The authors show that these are of order $O(\varepsilon^{-\alpha})$ and, after a suitable scaling, that they converge in distribution to an exponentially distributed

random variable. An important groundwork for this was laid in [ImPa1] although in this paper the behavior near the saddle points of U was not yet considered. As it turns out the lack of driving (deterministic) forces near these points leads to a delayed motion of X^ε. Since the Markov chain Z^ε is an approximation of X^ε one can expect a similar metastable behavior. Indeed, the main focus of Chapter 3 is on the proof of the following theorem:

Theorem 3.0.1. *There exists a constant $c(\alpha, h) > 0$ such that for every $t > 0$ and all $1 \leq i, j \leq n$ such that $i \neq j$ one has*

$$\lim_{\varepsilon \to 0} \mathbb{P}_x \left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t \rceil} = \mathfrak{m}_j \right) = \mathbb{P}_{\mathfrak{m}_i}(Y_t = \mathfrak{m}_j)$$

uniformly in $x \in \mathcal{S}$.

We proceed in a similar fashion to [ImPa2] and investigate first the exit times of a single potential well. We derive an analogical result: That is we show that these times as well as their mean values are of order $O(\varepsilon^{-\alpha})$. Due to the fact that, by definition, the local maxima are not contained in the state space $\mathcal{S}$ we do not have to deal with the dynamics near the saddle points of U. Afterwards we study the mean transition times between two fixed potential wells and show that they also increase polynomially in $\varepsilon^{-\alpha}$.

Let us remark that metastable behavior for discrete-time Markov chains in a more abstract context was studied in [BEGK2] and [Eck], among others. The main difference to this thesis is, on the one hand, the additional assumption of reversibility in [BEGK2], which is not fulfilled in our case, and on the other hand, the assumption that the potential wells are distinguishable in terms of their respective "attraction" , which can be expressed in terms of mean transition times. The latter assumption is fulfilled, for example, in the Gaussian case. In our case, however, it is not given since the mean transition times all have the same polynomial order $O(\varepsilon^{-\alpha})$.

In the fourth chapter we analyze the spectrum of the infinitesimal generator (or the discrete analog to be more precise) of the Markov chain Z^ε. Let us first take a look again at equation (0.0.1) with a Gaussian noise. Here the infinitesimal generator of the solution

has for sufficiently smooth test functions the form

$$(\boldsymbol{\mathcal{D}}^\varepsilon f)(x) = -U'(x)\frac{d}{dx}f(x) + \frac{\varepsilon^2}{2}\frac{d^2}{dx^2}f(x)$$

and is self-adjoint on the weighted L_2-space $L_2(\mathrm{e}^{-\frac{2U(x)}{\varepsilon^2}}\,dx)$. The spectral behavior of this generator has been studied for decades, see for example [BGK], [BuMa1], [BuMa2], [FrWe] and [KoMa], among others and can be summarized as follows: There are precisely n distinguished eigenvalues $\lambda_1^\varepsilon, \dots, \lambda_n^\varepsilon$ such that $\lambda_1^\varepsilon = 0$ and the other eigenvalues λ_i^ε, $2 \leq i \leq n$, are of exponentially small order, $\lambda_i^\varepsilon = O\left(\mathrm{e}^{-c_i/\varepsilon^2}\right)$, and separated from the remaining spectrum by a spectral gap. The values c_i again depend on the depths of the wells. That implies that if U is a generic potential, then one can "separate" the eigenvalues by their logarithmic rates. An analogical discrete situation is studied in [BEGK2]. Moreover, one can express these eigenvalues in terms of the inverse mean exit times from the potential wells. Indeed, Kolokoltsov and Makarov studied in [KoMa] (see also [Ko2], Chapter 8, Section 2) two different situations: Let δ^j be the indicator function of the well Ω_j. Then they considered in case A a compact interval M that contains all local minima of U as well as the inhomogeneous system

$$\boldsymbol{\mathcal{D}}^\varepsilon \tau^{j,\varepsilon} = \delta^j, \quad \tau^{j,\varepsilon}|_{M^c} = 0, \quad 1 \leq j \leq n.$$

In case B they considered the problem on the whole real line. Here the system has the form

$$\boldsymbol{\mathcal{D}}^\varepsilon \tau^{j,\varepsilon} = \delta^j - \pi^\varepsilon(\Omega_j)\mathbf{1}, \quad 1 \leq j \leq n, \tag{0.0.2}$$

where π^ε denotes the stationary distribution of X^ε and $\mathbf{1}(x) = 1$ for all $x \in \mathbb{R}$. It is well know that, in case A, the solution $\tau^{j,\varepsilon}(x)$ gives the time the process spent in Ω_j when started in x. Now let $\mathbf{G}^\varepsilon = (g_{i,j}^\varepsilon)_{i,j=1}^n$ be the matrix that consists of the averages (with respect to π^ε) of the solutions $\tau^{j,\varepsilon}$ over Ω_i, i.e.

$$g_{i,j}^\varepsilon := \frac{1}{\pi^\varepsilon(\Omega_i)}\int_{\Omega_i} \tau^{j,\varepsilon}(x)\pi^\varepsilon(dx).$$

If λ_i^ε, $1 \leq i \leq n$, denote the low lying eigenvalues of $\mathcal{D}^\varepsilon$ and μ_i^ε are the eigenvalues of $\mathbf{G}^\varepsilon$, then Theorem 1.2 in [KoMa] states that one has $\mu_i^\varepsilon = (\lambda_i^\varepsilon)^{-1}(1 + O(\varepsilon^\infty))$ (in case B it also

holds that $\mu_1^\varepsilon = \lambda_1^\varepsilon = 0$).

Furthermore, one knows that the corresponding eigenfunctions and eigenvectors, respectively, are almost constant within the different domains of attraction Ω_i.

If the noise term in equation (0.0.1) is given by a symmetric α-stable Lévy process, then the infinitesimal generator is a non-symmetric integro-differential operator of the form

$$(\mathcal{A}^\varepsilon f)(x) = -U'(x)\frac{d}{dx}f(x) + c(\alpha)\varepsilon^\alpha \int_{\mathbb{R}\setminus\{0\}} \frac{f(x+y) - f(x) - yf'(x)\mathbf{1}_{\{|y|\leq 1\}}(y)}{|y|^{\alpha+1}}dy.$$

where $c(\alpha)$ is a certain constant. The conjecture is that this operator possesses a similar spectral behavior, i.e. there is again a set of n eigenvalues $\lambda_1^\varepsilon = 0, \ldots, \lambda_n^\varepsilon$, that are separated from the remaining spectrum by a spectral gap. The close connection to mean exit times as well as the fact that these are of order $O(\varepsilon^{-\alpha})$ (see [ImPa1], [ImPa2]), lead us to the conclusion that these eigenvalues should be of polynomially small order $O(\varepsilon^\alpha)$. However, the spectral analysis of this operator turned out to be very difficult and for that reason we restrict ourselves in this thesis on the analysis of the spectrum of the generator of the Markov chain Z^ε which is viewed as a discrete approximation of X^ε. Therefore, the aforementioned heuristics stay the same which causes problems. More precisely, in the Gaussian case the different logarithmic rates allow to distinguish these eigenvalues although they are all exponentially small. In the analog discrete situation studied in [BEGK2] and [Eck] this is given by the notion of non-degeneracy which is expressed in terms of mean transition times which in turn are closely related to the eigenvalues. But since we expect the eigenvalues to have the same polynomially small order we are forced to use different methods. The approach for our analysis consists mainly in using the knowledge of Chapter 3 about the "correct" time scales. For finite, continuous-time Markov chains the relationship between the family $(\mathbf{P}_t)_{t\geq 0}$ of transition matrices and the infinitesimal generator $\mathbf{Q}$ is given by

$$\mathbf{P}_t = \text{“ e}^{t\mathbf{Q}}\text{ ”} = \sum_{k=0}^{\infty} \frac{t^k \mathbf{Q}^k}{k!}.$$

In the discrete-time case we have the equation $\mathbf{Q} = \mathbf{P} - \mathbf{I}$. By Theorem 3.0.1 we know that Z^ε behaves in the limit $\varepsilon \to 0$ statistically like a Markov jump process with the generator $\mathbf{Q}$ if we stretch the time by the factor $\frac{c(\alpha,h)}{\varepsilon^\alpha}$. Therefore, the conjecture is that

a part of the spectrum of the matrix $\mathbf{Q}^\varepsilon = \frac{c(\alpha,h)}{\varepsilon^\alpha}(\mathbf{P}^\varepsilon - \mathbf{I})$ converges to the spectrum $\sigma(\mathbf{Q}) = \left\{0, \lambda_2^{\mathbf{Q}}, \ldots, \lambda_n^{\mathbf{Q}}\right\}$ of $\mathbf{Q}$. Indeed, the following theorem holds:

Theorem 4.2.2. *The spectrum $\sigma(\mathbf{Q}^\varepsilon)$ can be divided into two disjoint parts $\sigma_1(\mathbf{Q}^\varepsilon)$ and $\sigma_2(\mathbf{Q}^\varepsilon)$ for which the following assertions hold:*

(i) $\sigma_1(\mathbf{Q}^\varepsilon)$ contains precisely n eigenvalues $\lambda_1^\varepsilon, \ldots, \lambda_n^\varepsilon$ such that $\lambda_1^\varepsilon = 0$ and

$$\lim_{\varepsilon \to 0} \lambda_i^\varepsilon = \lambda_i^{\mathbf{Q}}, \quad 2 \leq i \leq n.$$

(ii) There is a spectral gap that separates $\sigma_1(\mathbf{Q}^\varepsilon)$ and $\sigma_2(\mathbf{Q}^\varepsilon)$. More precisely, the following limit holds for the remaining eigenvalues $\lambda_{n+1}^\varepsilon, \ldots, \lambda_N^\varepsilon$:

$$\lim_{\varepsilon \to 0} |\lambda_i^\varepsilon| = \infty, \quad n+1 \leq i \leq N.$$

The most important step in the proof of Theorem 4.2.2 is the investigation of the asymptotic behavior of the characteristic polynomial of $\mathbf{Q}^\varepsilon$ and in regard to that we prove the following proposition:

Proposition 4.2.3. *Let P^ε and $P^{\mathbf{Q}}$ denote the characteristic polynomials of $\mathbf{Q}^\varepsilon$ and $\mathbf{Q}$, respectively. Then there are $K \geq 1$ and constants $c(\alpha, h) > 0$ and $C > 0$, the latter may also depend on the discretization parameters as well as α, such that for every $0 < \varepsilon < \varepsilon_0$ and all $\lambda \in \mathbb{C}$*

$$\left| \left(-\frac{\varepsilon^\alpha}{c(\alpha,h)} \right)^{N-n} P^\varepsilon(\lambda) - P^{\mathbf{Q}}(\lambda) \right| \leq C|\lambda|^K \varepsilon^\alpha.$$

In Section 4.3 we devote ourselves to the study of the eigenvectors and we confirm our conjecture that these are approximately constant within the domains of attraction of U. Let $\boldsymbol{\delta}^j = (\delta_x^j)_{x \in \mathcal{S}}$ be the indicator vector of $\mathcal{S} \cap [\mathfrak{s}_{j-1}, \mathfrak{s}_j]$, i.e. $\delta_x^j = 1$ if $x \in \mathcal{S} \cap [\mathfrak{s}_{j-1}, \mathfrak{s}_j]$ and 0 otherwise. Assume additionally that, for all $1 \leq i \leq n$, the eigenspace corresponding to $\lambda_i^{\mathbf{Q}}$ is one dimensional. Then:

Theorem 4.3.1. *Let $\lambda_1^\varepsilon, \ldots, \lambda_n^\varepsilon$ be the n eigenvalues from Theorem 4.2.2 (i) and $\boldsymbol{\psi}^{\varepsilon,i} = (\psi_x^{\varepsilon,i})_{x \in \mathcal{S}}$ the corresponding right eigenvectors normalized such that $\max_{1 \leq j \leq n} \left| \psi_{\mathfrak{m}_j}^{\varepsilon,i} \right| = 1$.*

Moreover, let $\boldsymbol{\psi}^{\mathbf{Q},i} = (\psi_j^{\mathbf{Q},i})_{j=1}^n$ *the right eigenvector of* $\mathbf{Q}$ *associated with the eigenvalue* $\lambda_i^{\mathbf{Q}}$ *normalized such that* $\max_{1\le j\le n}\left|\psi_j^{\mathbf{Q},i}\right| = 1$.. *Then we have* $\boldsymbol{\psi}^{\varepsilon,1} = (1,\dots,1)^T$ *and*

$$\lim_{\varepsilon\to 0}\left|\psi_x^{\varepsilon,i} - \sum_{j=1}^n \psi_{\mathfrak{m}_j}^{\mathbf{Q},i}\delta_x^j\right| = 0, \quad x\in\mathcal{S}, \quad 2\le i\le n.$$

For the derivation of this result we follow closely the ideas presented in [BEGK2] and [Eck].

Building on the results from Chapter 3 and 4 we conclude this thesis in Chapter 5 by a more detailed investigation of the relation between the eigenvalues from Theorem 4.2.2 (i) and mean exit times. This approach is motivated by [KoMa] and [Ko2], Chapter 8, in which, as mentioned before, the Gaussian case is studied. The discrete analog of the differential equations (0.0.2) are the following systems of inhomogeneous linear equations

$$(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\boldsymbol{\tau}^{\varepsilon,j} = \boldsymbol{\pi}^\varepsilon(\mathcal{S}_j)\mathbf{1} - \boldsymbol{\delta}^j, \quad 1\le j\le n.$$

Using the solutions $\boldsymbol{\tau}^{\varepsilon,j}$ of these systems we define a matrix $\mathbf{G}^\varepsilon = (g_{i,j}^\varepsilon)_{i,j=1}^n$ by

$$g_{i,j}^\varepsilon := \frac{\sum_{x\in\mathcal{S}_i}\tau_x^{\varepsilon,j}\pi_x^\varepsilon}{\sum_{x\in\mathcal{S}_i}\pi_x^\varepsilon}.$$

Then, given an additional and purely technical condition (B)[4], one can derive the following theorem.

Theorem 5.0.1. *Let* $\lambda_1^\varepsilon,\dots,\lambda_n^\varepsilon$ *be the eigenvalues of* $\mathbf{Q}^\varepsilon$ *from Theorem 4.2.2 (i) and let* $\mu_1^\varepsilon,\dots,\mu_n^\varepsilon$ *be the eigenvalues of* $\overline{\mathbf{G}}^\varepsilon := \frac{\varepsilon^\alpha}{c(\alpha,h)}\mathbf{G}^\varepsilon$. *Moreover, assume that condition (B) holds. Then*

$$\mu_1^\varepsilon = \lambda_1^\varepsilon = 0 \quad \textit{and} \quad \lim_{\varepsilon\to 0}\left|(\lambda_i^\varepsilon)^{-1} - \mu_i^\varepsilon\right| = 0, \quad 2\le i\le n.$$

In contrast to [KoMa] we are not able to prove this theorem "directly". Fortunately it turns out that one can construct solutions to these systems with help of the so called *fundamental matrix* and *deviation matrix*, respectively. These quantities were first introduced by Kemeny and Snell in [KeSn1] and [KeSn2]. For discrete-time Markov chains they are defined by

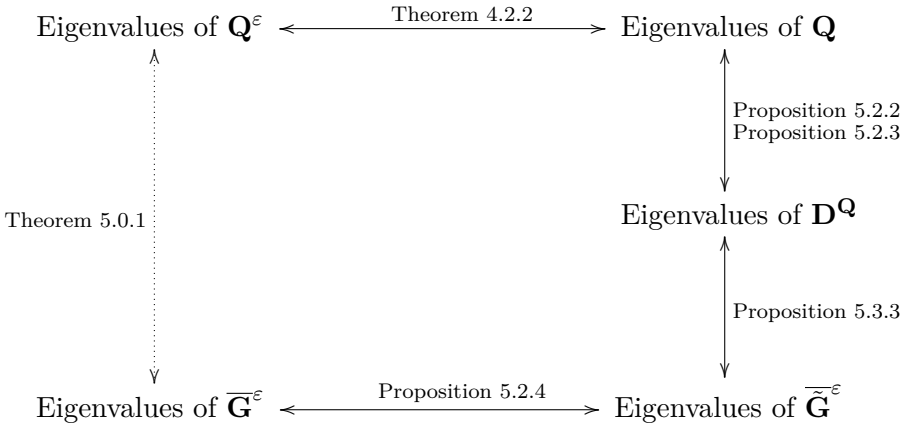
$$\mathbf{F}^\varepsilon := (\mathbf{I} - (\mathbf{P}^\varepsilon - \mathbf{\Pi}^\varepsilon))^{-1}$$

[4]The precise formulation of this condition is given in Chapter 5.

where $\mathbf{\Pi}^\varepsilon := (\pi_y^\varepsilon)_{x,y\in\mathcal{S}}$. The deviation matrix is then given by $\mathbf{D}^\varepsilon := \mathbf{F}^\varepsilon - \mathbf{\Pi}^\varepsilon$. The latter is the foundation for the matrix $\tilde{\mathbf{G}}^\varepsilon = (\tilde{g}_{i,j}^\varepsilon)_{i,j=1}^n$:

$$\tilde{g}_{i,j}^\varepsilon := \frac{\sum_{x\in\mathcal{S}_i} \pi_x^\varepsilon \left(\sum_{y\in\mathcal{S}_j} d_{x,y}^\varepsilon\right)}{\pi^\varepsilon(\mathcal{S}_j)}.$$

We show in Proposition 5.2.4 that this matrix has essentially the same eigenvalues as $\mathbf{G}^\varepsilon$. Let us denote by $\mathbf{D}^{\mathbf{Q}}$ the deviation matrix of the Markov chain Y and define $\overline{\tilde{\mathbf{G}}}^\varepsilon := \frac{\varepsilon^\alpha}{c(\alpha,h)}\tilde{\mathbf{G}}^\varepsilon$. Then the connection between the eigenvalues of all these matrices, and therefore the structure of the proof of Theorem 5.0.1, can be visualized with the following scheme:



Note that in the proof of Proposition 5.3.2 we use the close connection between the deviation matrix and the matrix $\mathbf{M} = (m_{x,y})_{x,y\in\mathcal{S}}$ of mean transition times on which we apply the results of Chapter 3.

1 Preliminaries

1.1 Notations

This section provides an overview over the notations we are going to use in this thesis. We start with some symbols. First, ":=" means "defined to be equal" while "$\stackrel{d}{=}$" denotes the equality in distribution. Second, $A \sqcup B$ denotes the union of two disjoint sets A and B.

We are going to use the Landau symbols in the usual way, i.e. for example

$$f^\varepsilon = O(g^\varepsilon) \Leftrightarrow \limsup_{\varepsilon \to 0} \left| \frac{f^\varepsilon}{g^\varepsilon} \right| < \infty \quad \text{as well as} \quad f^\varepsilon = o(g^\varepsilon) \Leftrightarrow \lim_{\varepsilon \to 0} \left| \frac{f^\varepsilon}{g^\varepsilon} \right| = 0.$$

Let $a \in \mathbb{C}$. The real part of a is denoted by $\mathrm{Re}(a)$ and the imaginary part is $\mathrm{Im}(a)$. Moreover, let $B_R(a)$ be the open ball centered around a with radius R, i.e.

$$B_R(a) \coloneqq \{z \in \mathbb{C} |\ |z - a| < R\}.$$

For linear operators on function spaces we use calligraphic letters such as $\mathcal{A}, \mathcal{D}$ and matrices will be denoted by capital bold letters such as $\mathbf{A}, \mathbf{B}$ etc. If the rows of a matrix $\mathbf{A}$ are indexed by a set $\mathcal{S}_1$ and the columns are indexed by set $\mathcal{S}_2$, then we write $\mathbf{A} = (a_{x,y})_{x \in \mathcal{S}_1}^{y \in \mathcal{S}_2}$ where $a_{x,y}$ denote the elements of $\mathbf{A}$. Sometimes it will be more convenient to write $(\mathbf{A})_{x,y}$ instead of $a_{x,y}$. In the special case $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S}$, i.e. $\mathbf{A}$ is a square matrix, we will write $\mathbf{A} = (a_{x,y})_{x,y \in \mathcal{S}}$. Also, for a given square matrix $\mathbf{A}$ we define by $\mathbf{A}_{\mathrm{dg}}$ the diagonal matrix where the entries on the main diagonal are $a_{x,x}$. Moreover, by $\mathbf{I}_K$ we denote the identity matrix of dimension $K \geq 1$ and by $\mathbf{E}_K$ the matrix in which every entry is 1. Note that we will omit the subscript K if there is no chance for confusion.

Vectors or, more generally, functions with a finite domain, say $\mathcal{S}$, will be denoted by small

bold letters, for example $\mathbf{f} = (f_x)_{x\in\mathcal{S}}$. Usually, vectors are given as column vectors and the transpose, e.g. $\mathbf{f}^T$, is a row vector. Furthermore, we write $\mathbf{1} \coloneqq (1, 1, \ldots, 1)^T$ and $\mathbf{0} \coloneqq (0, \ldots, 0)^T$.
The standard inner product for two vectors $\mathbf{u}, \mathbf{v}$ is denoted by $\langle \mathbf{u}, \mathbf{v}\rangle \coloneqq \sum_{x\in\mathcal{S}} u_x v_x$. Also, for another given vector $\boldsymbol{\mu}$ we define the weighted inner product by $\langle \mathbf{u}, \mathbf{v}\rangle_{\boldsymbol{\mu}} \coloneqq \sum_{x\in\mathcal{S}} u_x v_x \mu_x$.
Norms induced by inner products are then denoted by $\|\cdot\|$ and $\|\cdot\|_{\boldsymbol{\mu}}$, respectively.

The symbol for a probability measure is $\mathbb{P}$ with corresponding expectation $\mathbb{E}$. The weak convergence of a sequence $(X_k)_{k\geq 1}$ to X is denoted by $X_k \Rightarrow X,\ k \to \infty$.
In the context of Markov chains it is common to use conditional probabilities and expectations and the condition is usually the initial value of the chain. This will be denoted by $\mathbb{P}_x$ and $\mathbb{E}_x$, respectively.
Finally, $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra.

1.2 Laplace Transforms of Random Variables

This section is devoted to the concept of Laplace transforms for random variables. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 1.2.1. Fix an event $A \in \mathcal{F}$ and let X be a real valued random variable. The *Laplace transform* of X with respect to A is defined by

$$\varphi_{X,A}(u) \coloneqq \mathbb{E}\left[\mathrm{e}^{uX}\,\mathbf{1}_A\right] \tag{1.2.1}$$

for all $u \in \mathbb{C}$ for which the expectation is finite.

Note that, although we use complex numbers as arguments, the question whether this expectation is finite or not only depends on the real part of u.
Unfortunately, in the literature the notation "Laplace transform" is not always used in the same way. First of all, mostly there is no restriction on a certain event A, i.e. Definition 1.4.1 is used with $A = \Omega$. In this case we will use the notation *standard Laplace*

transform. The most common definition is probably the case $A = \Omega$ and $u \leq 0$ since the Laplace transform was initially defined as a more convenient tool (compared to, for example, characteristic functions) to investigate random variables with values in $[0, \infty)$. A good reference for this case are the books [Ka] and [FrGr]. If one additionally allows positive values for u, then the Laplace transform is often referred to as *moment generating function*. We will justify this name later in this section. However, inspired by the papers [BEGK2] and [Eck] we will define Laplace transforms as in (1.2.1).
Let us start by stating the uniqueness and continuity theorem for standard Laplace transforms of random variables with values in $[0, \infty)$.

Proposition 1.2.2. *Let $X, Y, X_k,\ k \geq 1$, be random variables with values in $[0, \infty)$. Then*

(i) If $\varphi_{X,\Omega}(u) = \varphi_{Y,\Omega}(u)$ for all $u \leq 0$, then $X \stackrel{d}{=} Y$.

(ii) If $\lim_{k\to\infty} \varphi_{X_k,\Omega}(u) = \varphi_{X,\Omega}(u)$ for all $u \leq 0$, then $X_k \Rightarrow X,\ k \to \infty$.

A proof can be found, for example, in [Ka], Theorem 5.3.

As mentioned, the following proposition shows how and when we can compute moments of a random variable X if we know its Laplace transform.

Proposition 1.2.3. *Fix $A \in \mathcal{F}$ and let X be a real valued random variable. Assume that $\varphi_{X,A}(u)$ is finite in a R-neighborhood of 0, i.e for some $R > 0$ and all $u \in B_R(0)$. Then, $\varphi_{X,A}$ is analytic in that neighborhood and hence can be expanded in a Taylor series.*
Moreover, in the case $A = \Omega$ this series is given by

$$\varphi_{X,\Omega}(u) = \sum_{k=0}^{\infty} \frac{u^k}{k!} \mathbb{E}(X^k)$$

and hence the moments of X are finite and can be computed with the formula

$$\mathbb{E}(X^k) = \varphi_{X,\Omega}^{(k)}(0), \quad k \geq 1,$$

where $\varphi_{X,\Omega}^{(k)}$ denotes the k-th derivative of $\varphi_{X,A}$.

We conclude this section with two important examples that will be used in this thesis.

Example: Fix $A = \Omega$.

(a) Let X be exponentially distributed with parameter $\beta > 0$. Then

$$\varphi_{X,\Omega}(u) = \frac{\beta}{\beta - u}, \quad u < \beta.$$

(b) Let X be geometrically distributed with parameter $p \in (0,1)$, i.e. $\mathbb{P}(X = k) = p(1-p)^{k-1},\ k \geq 1$. Then

$$\varphi_{X,\Omega}(u) = \frac{p\,\mathrm{e}^u}{1-(1-p)\,\mathrm{e}^u}, \quad u < -\ln(1-p).$$

1.3 Markov Processes

In this section we want to give a very brief overview over the topics from the field of Markov processes that we are going to need in later chapters of this thesis. Obviously there is a huge amount of literature on Markov processes available. As the basis for what follows we will use the books [App], [EtKu], [FrGr], [Ka], [No] and [Li] where the interested reader can also find the proofs of the results below.

1.3.1 General Definition

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $T = \mathbb{N}_0$ or $T = [0, \infty)$ an index set. Let $(\mathcal{F}_t)_{t\in T}$ be a filtration, i.e. an increasing sequence of sub-σ-algebras $\mathcal{F}_t \subseteq \mathcal{F}$. Recall that a stochastic process $X = (X_t)_{t\in T}$ is called *adapted* to this filtration if X_t is $\mathcal{F}_t$-measurable for every $t \in T$. The collection $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathbb{P})$ is called filtered probability space.
The most general definition of an $\mathbb{R}$-valued Markov process is the following.

Definition 1.3.1. An $\mathbb{R}$-valued stochastic process $X = (X_t)_{t\in T}$ adapted to the filtration $(\mathcal{F}_t)_{t\in T}$ is called *Markov process* if

$$\mathbb{P}(X_{t+s} \in B |\, \mathcal{F}_t) = \mathbb{P}(X_{t+s} \in B |\, X_t) \quad a.s. \tag{1.3.1}$$

for all $s, t \in T$ and all $B \in \mathcal{B}(\mathbb{R})$. Formula (1.3.1) is called the *Markov property*.

For our purposes it is sufficient to restrict ourselves to the very important subclass of (time-)homogeneous Markov processes.

Definition 1.3.2. Define a family of *transition probabilities* $p_{t,x}(B)$, $x \in \mathbb{R}$, $t \in T$, $B \in \mathcal{B}(\mathbb{R})$, by

$$p_{t,x}(B) \coloneqq \mathbb{P}(X_t \in B |\, X_0 = x). \tag{1.3.2}$$

Then, a Markov process $X = (X_t)_{t\in T}$ is said to be *(time-)homogeneous* if equation (1.3.1) can be written as

$$\mathbb{P}(X_{t+s} \in B |\, \mathcal{F}_t) = p_{t+s,X_t}(B) \quad a.s. \tag{1.3.3}$$

for all $s, t \in T$ and $B \in \mathcal{B}(\mathbb{R})$.

Let us introduce the so called strong Markov property. For this recall that a *stopping time* $\tau \colon \Omega \to [0, \infty]$ is random variable for which $\{\tau \leq t\} \in \mathcal{F}_t$ holds for each $t \geq 0$.

Definition 1.3.3 Let $X = (X_t)_{t\in 0}$ be adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$. Denote by $p_{s,x}$ its transition probabilities and assume that the mapping $x \mapsto p_{s,x}(B)$ is measurable for every $s \in T$ and $B \in \mathcal{B}(\mathbb{R})$. Moreover, let τ be an almost surely finite stopping time with respect to this filtration. Then X is said to have the *strong Markov property* if

$$\mathbb{P}(X_{\tau+s} \in B |\, \mathcal{F}_\tau) = p_{\tau+s,X_\tau}(B) \quad a.s.$$

for all $s \in T$ and $B \in \mathcal{B}(\mathbb{R})$.

Generally, the strong Markov property is more restrictive than the simple Markov property, i.e. there are Markov processes that do not have the strong Markov property, see [FrGr] for a counterexample. However, in the cases that are of interest for us the strong Markov property is always ensured.

Proposition 1.3.4. *Let $X = (X_t)_{t\in T}$ be a homogeneous Markov process. If $T = \mathbb{N}_0$ or if the state space of X is finite, then X has the strong Markov property.*

For a proof see Theorem 1.4.2 and Theorem 2.8.1 in [No].

Let us now consider these special cases in more detail. From now on assume that the filtration is the natural filtration, i.e. $\mathcal{F}_s = \sigma(X_u, u \leq s)$. Also, let X take values in a finite subset $\mathcal{S} \subset \mathbb{R}$. In this case X is called *Markov chain* with the state space $\mathcal{S}$.

1.3.2 Discrete-time Markov Chains on a Finite State Space

For this section let $T = \mathbb{N}_0$. Then Definition 1.3.1 can be simplified.

Definition 1.3.5. A sequence $X = (X_k)_{k\geq 0}$ with values in a finite set $\mathcal{S}$ is called a Markov chain if for all $k \geq 0$ and all collections of states $x_0, \ldots, x_{k+1} \in \mathcal{S}$

$$\mathbb{P}(X_{k+1} = x_{k+1} |\, X_k = x_k, \ldots, X_0 = x_0) = \mathbb{P}(X_{k+1} = x_{k+1} |\, X_k = x_k). \tag{1.3.4}$$

Formula (1.3.4) implies that a homogeneous Markov chain $X = (X_k)_{k\geq 0}$ on a finite state space $\mathcal{S}$ is uniquely characterized by an initial distribution $\boldsymbol{\mu} = (\mu_x)_{x\in\mathcal{S}}$ given by

$$\mu_x \coloneqq \mathbb{P}(X_0 = x)$$

and a matrix $\mathbf{P} = (p_{x,y})_{x,y\in\mathcal{S}}$. The latter is called *one step transition matrix* and defined by

$$p_{x,y} \coloneqq \mathbb{P}(X_1 = y |\, X_0 = x) = \mathbb{P}_x(X_1 = y).$$

Indeed, for every $k \geq 1$ and every $x \in \mathcal{S}$ we have

$$\mathbb{P}(X_k = x) = (\boldsymbol{\mu}^T \mathbf{P}^k)_x.$$

The matrix $\mathbf{P}^k = \left(p^{(k)}_{x,y}\right)_{x,y\in\mathcal{S}}$ with $p^{(k)}_{x,y} = \mathbb{P}(X_k = y |\, X_0 = x)$ is called the k-th step transition matrix. Note that the connection between the one step transition matrix and the transition probabilities from Definition 1.3.2 is given by

$$p_{x,y} = p_{1,x}(\{y\}).$$

The following proposition states some obvious properties of a transition matrix.

Proposition 1.3.6. *Let $\mathbf{P} = (p_{x,y})_{x,y\in Ss}$ be the one step transition matrix of a Markov chain $X = (X_k)_{k\geq 0}$. Then*

(i) $0 \leq p_{x,y} \leq 1$ *for all* $x, y \in \mathcal{S}$.

(ii) $\sum_{y \in \mathcal{S}} p_{x,y} = 1$ *for all* $x \in \mathcal{S}$.

Now let us come to the the notion of a stationary (or invariant) distribution.

Definition 1.3.7. A *stationary distribution* $\boldsymbol{\pi} = (\pi_x)_{x \in \mathcal{S}}$ for a discrete-time Markov chain with transition matrix $\mathbf{P}$ is a vector that solves the following linear system of equations:

$$\boldsymbol{\pi}^T \mathbf{P} = \boldsymbol{\pi}^T, \qquad \sum_{x \in \mathcal{S}} \pi_x = 1, \quad \pi_x \geq 0, \ x \in \mathcal{S}. \tag{1.3.5}$$

Note that this system always possesses a solution. Indeed, since $\mathbf{1}$ is obviously a right eigenvector of $\mathbf{P}$ corresponding to the eigenvalue 1 the eigenspace is at least one-dimensional. Hence there exists a left eigenvector that can be normalized such that (1.3.5) is fulfilled. The question is then whether this solution is unique or not or, equivalently, when the eigenspace corresponding to the eigenvalue 1 is one-dimensional. This problem is closely related to the notions of irreducibility, recurrence and transience.

Definition 1.3.8 Let $X = (X_k)_{k \geq 0}$ a discrete-time Markov chain with state space $\mathcal{S}$.

(i) X is called *irreducible* if all states communicate with each other, i.e. if for any $x, y \in \mathcal{S}$ there is $k \geq 0$ such that

$$p_{x,y}^{(k)} > 0.$$

(ii) A state $x \in \mathcal{S}$ is called *recurrent* if

$$\mathbb{P}_x(X_k = x \text{ for infinitely many } k) = 1.$$

The chain X is called recurrent if every state is recurrent.

(iii) A state $x \in \mathcal{S}$ is called *transient* if

$$\mathbb{P}_x(X_k = x \text{ for infinitely many } k) = 0.$$

Irreducibility means that the Markov chain will reach every state y with positive probability from every initial state x after finitely many steps. A recurrent state is a state to which the Markov chain keeps coming back while a transient state is one that the chain will eventually leave forever.
The following proposition provides a sufficient condition for the solution of (1.3.5) to be unique and is proven in [KeSn1], Theorem 4.1.4.

Proposition 1.3.9. *Let* $\mathbf{P}$ *be the one step transition matrix for a homogeneous Markov chain* X*. If, for some* $K \geq 1$, $\mathbf{P}^K$ *has non-zero entries, then the linear system (1.3.5) has a unique solution.*

Note that if there exists some $K \geq 1$ such that $\mathbf{P}^K$ has non-zero entries, the Markov chain is called *regular*.

1.3.3 Continuous-time Markov Chains on a Finite State Space

Now let $T = [0, \infty)$ and $\mathcal{S}$ be a finite set. Similarly to the discrete-time case, homogeneous Markov chains with a continuous time parameter are uniquely characterized by two quantities. Here it is again an initial distribution and a so called Q-matrix.
Definition 1.3.10. A matrix $\mathbf{Q} = (q_{x,y})_{x,y\in\mathcal{S}}$ is called a *Q-matrix* if it has the following properties:

(i) $q_{x,y} \geq 0$ for all $x, y \in \mathcal{S}$ such that $x \neq y$.

(ii) $\sum_{y\in\mathcal{S}} q_{x,y} = 0$ for all $x \in \mathcal{S}$.

Now let $X = (X_t)_{t\geq 0}$ be a continuous-time homogeneous Markov chain with the corresponding family of transition matrices $\mathbf{P}(t) = (p_{x,y}(t))_{x,y\in\mathcal{S}}$ given by $p_{x,y}(t) = \mathbb{P}(X_t = y|\ X_0 = x)$. Since $\mathcal{S}$ is finite there is a one-to-one correspondence between $\mathbf{P}(t)$ and Q-matrices. Indeed, consider the matrix $\mathbf{Q} = (q_{x,y})_{x,y\in\mathcal{S}}$ given by

$$q_{x,y} = \frac{d}{dt}p_{x,y}|_{t=0} \quad \text{if } x \neq y \quad \text{and} \quad -q_{x,x} = \frac{d}{dt}p_{x,x}|_{t=0}.$$

Then this matrix $\mathbf{Q}$ is a Q-matrix.
On the other hand, if a Q-matrix $\mathbf{Q}$ is given one can construct a family $\mathbf{P}(t)$ of transition

matrices by setting

$$\mathbf{P}(t) := \text{“ } \mathrm{e}^{t\mathbf{Q}} \text{ ”} := \sum_{k=0}^{\infty} \frac{t^k \mathbf{Q}^k}{k!}.$$

Hence, one can say that a Q-matrix generates a Markov chain and is therefore called *infinitesimal generator*. We will elaborate on that topic in the next section.

The entries of a generator matrix have a probabilistic meaning. Indeed, if a chain X with generator $\mathbf{Q}$ starts in a state $x \in \mathcal{S}$, then it stays there for an $\mathrm{Exp}(-q_{x,x})$-distributed time before it jumps to another state y with probability $\frac{q_{x,y}}{-q_{x,x}}$. Generally, a continuous-time Markov chain is uniquely characterized by the sequence of these jump times and the sequence of states after these jumps. The latter sequence is a discrete-time Markov chain itself and is called *embedded jump chain*.

Similar to the last section let us conclude with the definition of a stationary distribution, this time in terms of the generator matrix $\mathbf{Q}$.

Definition 1.3.11. A *stationary distribution* $\boldsymbol{\pi} = (\pi_x)_{x\in\mathcal{S}}$ for a continuous-time Markov chain with generator matrix $\mathbf{Q}$ is a vector that solves the following linear system of equations:

$$\boldsymbol{\pi}^T \mathbf{Q} = \mathbf{0}^T, \qquad \sum_{x\in\mathcal{S}} \pi_x = 1, \quad \pi_x \geq 0,\ x \in \mathcal{S}. \tag{1.3.6}$$

From Definition 1.3.10 (ii) we see that 0 is an eigenvalue of $\mathbf{Q}$ corresponding to the right eigenvector $\mathbf{1}$ and therefore the system (1.3.6) always possesses a solution. The question of uniqueness can be answered by checking if the embedded jump chain is irreducible and recurrent. This leads to the following sufficient condition which is proven in [No], Theorem 3.5.3.

Proposition 1.3.12. *Let $X = (X_t)_{t\geq 0}$ be a continuous-time Markov chain with generator matrix $\mathbf{Q}$. If $|q_{x,y}| > 0$ for all $x, y \in \mathcal{S}$, then X possesses a unique stationary distribution.*

1.3.4 Transition Semigroups and Infinitesimal Generators

As we mentioned in the last section a Q-matrix is the infinitesimal generator for a homogeneous continuous-time Markov chain. In this section we want to give the theoretical

background for these objects.

Let $X = (X_t)_{t\geq 0}$ be an $\mathbb{R}$-valued homogeneous Markov process. First we define a family of linear operators called *transition semigroup*. Denote

$$B_b(\mathbb{R}) \coloneqq \{f : \mathbb{R} \to \mathbb{R} |\ f \text{ bounded and measurable}\}$$

and recall that this space equipped with the supremum-norm $\|f\|_\infty \coloneqq \sup_{x\in\mathbb{R}} |f(x)|$ is a Banach space.

Definition 1.3.13. The *transition semigroup* of X is the family of linear operators $(\mathcal{T}_t)_{t\geq 0} : B_b(\mathbb{R}) \to B_b(\mathbb{R})$ defined by

$$(\mathcal{T}_t f)(x) \coloneqq \mathbb{E}_x f(X_t).$$

We can now come to the definition of the infinitesimal generator of this semigroup.

Definition 1.3.14. Let $(\mathcal{T}_t)_{t\geq 0}$ be the transition semigroup of a homogeneous Markov process $X = (X_t)_{t\geq 0}$. The *infinitesimal generator* of this semigroup is defined by

$$\mathcal{A}f \coloneqq \lim_{t\to 0} \frac{\mathcal{T}_t f - f}{t}$$

where this limit is taken in $B_b(\mathbb{R})$. The domain $\mathrm{dom}(\mathcal{A})$ of this operators is the collection of all functions for which this limit exists.

Example. Let X be a standard Brownian motion. Then, for sufficiently smooth functions, $\mathcal{A} = \frac{1}{2}\boldsymbol{\Delta}$ where $\boldsymbol{\Delta}$ is the Laplace operator.

Obviously, if we consider a finite state space $\mathcal{S}$, then these concepts simplify significantly. We already mentioned the case of a continuous-time Markov chain. Let us state that in a proposition.

Proposition 1.3.15. *Let $X = (X_t)_{t\geq 0}$ be a homogeneous Markov chain with generator matrix $\mathbf{Q}$. Then $\mathbf{Q}$ is the infinitesimal generator of X in the sense of Definition 1.3.14.*

In the case of a discrete-time Markov chain $X = (X_k)_{k\geq 0}$ with one step transition matrix $\mathbf{P}$ one actually does not need "infinitesimal characteristics". However, one can show that

$$\mathbf{P} = \sum_{k=0}^{\infty} \frac{(\mathbf{P}-\mathbf{I})^k}{k!} = \mathrm{e}^{\mathbf{P}-\mathbf{I}}.$$

Therefore, the matrix $\mathbf{P}-\mathbf{I}$ can be viewed as the discrete-time analog of a generator matrix $\mathbf{Q}$. Indeed, one can easily verify that $\mathbf{P}-\mathbf{I}$ is a Q-matrix in the sense of Definition 1.3.10. This is also consistent with the definition of a stationary distribution since we can rewrite Definition 1.3.7 to

$$\boldsymbol{\pi}^T(\mathbf{P}-\mathbf{I}) = \mathbf{0}^T, \qquad \sum_{x\in\mathcal{S}} \pi_x = 1, \quad \pi_x \geq 0,\ x \in \mathcal{S}.$$

1.4 α-stable Lévy Processes

Let again $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

In this section we provide the main definitions and properties regarding symmetric α-stable Lévy processes which are the class of stochastic processes that play the leading role in this thesis. We also state the existence and uniqueness results regarding Lévy-driven stochastic differential equations and give a formula for the infinitesimal generator. The definitions and statements are mainly taken from [App] and [Sa].

1.4.1 Definition and Properties

We start with the introduction of stable random variables.

Definition 1.4.1 A random variable X is called *stable* if there exist sequences $(c_n)_{n\geq 1}$ and $(d_n)_{n\geq 1}$ with $c_n > 0$ such that for each $n \geq 1$

$$X_1 + X_2 + \ldots + X_n \stackrel{d}{=} c_n X + d_n,$$

where $X_1, \ldots, X_n$ are independent copies of X. If $d_n = 0$, then X is called *strictly stable*. We call a probability distribution stable if it is the law of a stable random variable.

It is known (see for example [Fell]) that from this definition it follows that $c_n = cn^{\frac{1}{\alpha}}$ for some $c > 0$ and $0 < \alpha \leq 2$. Therefore, stable random variables are often referred to as

α-stable and the number α is called *index of stability.*
Moreover, we call a random variable *symmetric* if it has a symmetric probability law.

There are multiple ways to introduce and parametrize stable distributions, see for example [Berg], [SaTa], [Zo] as well as [Ko1] or [Ko2], Chapter 5, for more details. The following proposition is taken from [Berg].

Proposition 1.4.1. *Let X be a symmetric α-stable random variable with $0 < \alpha < 2$ and let p_α denote its probability density function. Then*

$$p_\alpha(x) = \sum_{k=1}^{n} \frac{c_k(\alpha)}{|x|^{\alpha k+1}} + O\left(\frac{1}{|x|^{\alpha(n+1)+1}}\right) \quad as \quad |x| \to \infty,$$

where

$$c_k(\alpha) = \frac{(-1)^{k+1}}{\pi k!}\Gamma(\alpha k+1)\sin\left[k\frac{\alpha\pi}{2}\right].$$

An immediate consequence of this result is the following corollary.

Corollary 1.4.2. *Let X be a symmetric α-stable random variable and let $0 < a < b$ be sufficiently large. Then there exists a constant $C = C(\alpha) > 0$ such that*

$$\left|\mathbb{P}(a \le X \le b) - \left(\frac{c_1(\alpha)}{\alpha a^\alpha} - \frac{c_1(\alpha)}{\alpha b^\alpha}\right)\right| \le C(\alpha)\left(\frac{1}{a^{2\alpha}} + \frac{1}{b^{2\alpha}}\right).$$

Now let us introduce Lévy processes. Since in this thesis we are going to deal with a certain subclass, namely symmetric α-stable processes, the following definition is slightly more specific than the usual one.

Definition 1.4.3. A stochastic process $L = (L_t)_{t\ge 0}$ is called symmetric α-stable Lévy process if it satisfies the following conditions:

(i) $L_0 = 0$ $\mathbb{P}$-a.s.,

(ii) L has independent and stationary increments,

(iii) L is stochastically continuous,

(iv) All finite dimensional distributions of L are symmetric and α-stable.

The properties (i)-(iii) are the ones that are usually used to define Lévy processes while the last one represents the restriction to the aforementioned subclass.
If we equip our probability space with a filtration such that the Lévy processes is adapted to it, for example the augmented filtration of a Lévy process L, then we can include these processes in the class of Markov processes.

Proposition 1.4.4. *Every Lévy process is a Markov process. Moreover, every Lévy process also has the strong Markov property.*

In view of this statement one may ask if it is possible to give an explicit form for the infinitesimal generator for a symmetric α-stable Lévy process. And indeed, at least for test functions from the Schwartz space $S(\mathbb{R})$ defined by

$$S(\mathbb{R}) \coloneqq \left\{ f \in C^\infty(\mathbb{R}) | \sup_{x\in\mathbb{R}} \left| x^m \frac{d^n}{dx^n} f(x) \right| < \infty,\ m,n \geq 0 \right\}$$

this is possible. For more details see [Ja].

Proposition 1.4.5. *Let L be a symmetric α-stable Lévy process with index $0 < \alpha < 2$ and denote by $\mathcal{A}$ its infinitesimal generator. Then*

$$(\mathcal{A}f)(x) = c(\alpha) \int_{\mathbb{R}\setminus\{0\}} \frac{f(x+y) - f(x) - yf'(x)\mathbf{1}_{\{|y|\leq 1\}}(y)}{|y|^{\alpha+1}} dy, \quad f \in S(\mathbb{R}),$$

where $c(\alpha)$ is a certain constant.

Note that this operator is often denoted by $-(-\boldsymbol{\Delta})^{\frac{\alpha}{2}}$ and is referred to as *fractional Laplacian*.
We conclude this overview by stating a property that is important for our discretization scheme and is known as *self-similarity*.

Proposition 1.4.6. *Let $a \geq 0$ and $L = (L_t)_{t\geq 0}$ be an α-stable Lévy process. Then*

$$a^{\frac{1}{\alpha}} L_t \overset{d}{=} L_{at} \quad \textit{for all } t \geq 0.$$

1.4.2 Lévy-driven Stochastic Differential Equations

The theory for stochastic differential equations is very rich. A good introduction to the general theory can be found in [Pr], for example. We will only state the results for the case that concerns us. They are taken from [App].
So let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space such that the filtration satisfies the so called *usual hypotheses*, i.e. it is complete and right-continuous. Consider the equation

$$X_t = x + \int_0^t b(X_{s-})ds + L_t, \tag{1.4.1}$$

where $x \in \mathbb{R}$, $X_{s-} := \lim_{u\uparrow s} X_u$ and $(L_t)_{t\geq 0}$ is a symmetric α-stable Lévy process.
First one should think about the question of existence and uniqueness of a solution. By "solution" we always mean a *strong solution* in the sense of [Pr], i.e. a stochastic process $X = (X_t)_{t\geq 0}$ adapted to the given filtration that has right-continuous path with left limits (i.e. càdlàg-paths) and that fulfills equation (1.4.1).

Proposition 1.4.7. *Let $b\colon \mathbb{R} \to \mathbb{R}$ be a globally Lipschitz-continuous function, i.e. there exists a constant $K > 0$ such that $|b(x)-b(y)| \leq K|x-y|$. Then, equation (1.4.1) possesses a unique strong solution.*

The next proposition states the Markov property of a solution.

Proposition 1.4.8. *The solution to (1.4.1) is a homogeneous Markov process.*

We finish this short excursion by looking at the infinitesimal generator of the solution.

Proposition 1.4.9. *Let $\mathcal{A}$ denote the infinitesimal generator of the solution $X = (X_t)_{t\geq 0}$ of equation (1.4.1). Then, $\mathcal{A}$ has the following representation on the space $C_0^2(\mathbb{R})$ of twice differentiable functions with compact support:*

$$(\mathcal{A}f)(x) = b'(x)f'(x) + c(\alpha)\int_{\mathbb{R}\setminus\{0\}} \frac{f(x+y) - f(x) - yf'(x)\mathbf{1}_{\{|y|\leq 1\}}(y)}{y^{1+\alpha}}dy$$

for some constant $c(\alpha)$.

2 From the SDE to the Markov Chain

From now on fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ where the filtration satisfies the usual hypotheses.

Consider the equation

$$X_t^\varepsilon(x) = x - \int_0^t U'(X_s^\varepsilon)ds + \varepsilon L_t. \tag{2.0.1}$$

Here $\varepsilon > 0$ is a small parameter and L is a symmetric α-stable Lévy process with $0 < \alpha < 2$. Furthermore, let us assume that the potential function U has the following properties:

(U1) U is continuously differentiable on the whole real line and three times continuously differentiable on a compact interval $[-K, K]$ for some $K > 0$ large enough.

(U2) $|U'(x)| > c_1|x|^{1+c_2}$ as $x \to \pm\infty$ for some $c_1, c_2 > 0$.

(U3) U has exactly n local minima $\mathfrak{m}_i$, $1 \leq i \leq n$, as well as $n-1$ local maxima $\mathfrak{s}_i$, $1 \leq i \leq n-1$, enumerated in increasing order

$$-\infty = \mathfrak{s}_0 < \mathfrak{m}_1 < \mathfrak{s}_1 < \mathfrak{m}_2 < \cdots < \mathfrak{m}_n < \mathfrak{s}_n = \infty. \tag{2.0.2}$$

Moreover, assume all these extrema are non-degenerate, i.e.

$$U''(\mathfrak{m}_i) > 0,\ 1 \leq i \leq n, \quad \text{and} \quad U''(\mathfrak{s}_i) < 0,\ 1 \leq i \leq n-1,$$

and denote

$$\mathcal{M} \coloneqq \{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}.$$

Since U does not fulfill the Lipschitz condition from Proposition 1.4.7 one must show the existence of a strong solution first. That is done in [ImPa2].

2.1 Time and Space State Discretization

First, we perform an Euler type discretization of the time line. For this purpose let $h > 0$ be the step size for the time. The Euler method applied to (2.0.1) yields

$$X^{\varepsilon}_{kh} = X^{\varepsilon}_{(k-1)h} - hU'(X^{\varepsilon}_{(k-1)h}) + \varepsilon(L_{kh} - L_{(k-1)h}).$$

Using Proposition 1.4.6 and the independent increments of the process L we derive

$$L_{kh} - L_{(k-1)h} \stackrel{d}{=} h^{\frac{1}{\alpha}} L_1.$$

Let $(\xi_k)_{k\geq 1}$ be a sequence of independent and identically distributed random variables such that $\xi_1 \stackrel{d}{=} L_1$. Let us define recursively a sequence $(\tilde{X}_k)_{k\geq 0}$ by

$$\tilde{X}^{\varepsilon}_0 \coloneqq x, \quad \tilde{X}^{\varepsilon}_{kh} \coloneqq \tilde{X}^{\varepsilon}_{(k-1)h} - hU'(\tilde{X}^{\varepsilon}_{(k-1)h}) + \varepsilon h^{\frac{1}{\alpha}} \xi_k, \quad k \geq 1. \tag{2.1.1}$$

Second, let $R \in \mathbb{N}$ such that $-R < \mathfrak{m}_1$ and $\mathfrak{m}_n < R$ an. We divide the interval $[-R, R]$ in the domains of attraction of U, i.e. $[-R, R] = \Omega_1 \sqcup \ldots \sqcup \Omega_n$ where

$$\Omega_1 \coloneqq [-R, \mathfrak{s}_1),\ \Omega_2 \coloneqq [\mathfrak{s}_1, \mathfrak{s}_2), \ldots,\ \Omega_n \coloneqq [\mathfrak{s}_{n-1}, R].$$

Now we can define a set $\mathcal{S}$ that will later serve as the state space for a Markov chain. For this fix a small parameter $\delta > 0$ and write each Ω_i as a disjoint union of intervals of the form $[a, b)$ (note that the interval on the right end of Ω_n is of the form $[a, b]$) such that the maximum length of these intervals does not exceed δ, i.e.

$$\Omega_i = \bigsqcup_{k=1}^{N(i,\delta)} I^{(i)}_k \qquad \text{and} \qquad \max_{1\leq i\leq n} \max_{1\leq k\leq N(i,\delta)} |I^{(i)}_k| < \delta.$$

Let us give the notations for the boundaries of the intervals $I^{(i)}_k$:

$$a^{(i)}_k \coloneqq \inf\{x |\ x \in I_k\}, \quad b^{(i)}_k \coloneqq \sup\{x |\ x \in I_k\}. \tag{2.1.2}$$

One should keep in mind that, even though it is suppressed in the notation above, the quantities $a^{(i)}_k$ and $b^{(i)}_k$ depend on the parameter δ. In fact, for every $\delta > 0$ one has

$$b^{(i)}_k - a^{(i)}_k \leq \delta.$$

Obviously the natural numbers $N(i,\delta)$ fulfill $\lim_{\delta\to 0} N(i,\delta) = \infty,\ 1 \le i \le n$.
With $N(\delta) \coloneqq N(1,\delta) + \ldots + N(n,\delta)$ and after renumbering the intervals I_k we can write

$$[-R,R] = \bigsqcup_{k=1}^{N} I_k \qquad \text{and} \qquad \Omega_i = \bigsqcup_{k=N(1,\delta)+\ldots+N(i-1,\delta)+1}^{N(1,\delta)+\ldots+N(i,\delta)} I_k. \tag{2.1.3}$$

To avoid technical difficulties we additionally assume that δ is chosen in such a way that the local minima do not coincide with the interval boundaries,

$$\mathcal{M} \cap \{a_1, b_1, \ldots, a_N, b_N\} = \emptyset. \tag{2.1.4}$$

Now we want to construct the state space for our Markov chain. This should contain all minima but no maximum of U. Hence, define a set $\mathcal{S}$ with N elements $x_1, \ldots, x_N$ in the following way:

(**S1**) $\mathcal{S}$ contains exactly one point out of the interior of each interval, i.e. for
$1 \le k \le N$ choose $x_k \in I_k \setminus \{a_k, b_k\}$.

(**S2**) If I_k contains a local minimum of U, say $\mathfrak{m}_i$, then set $x_k = \mathfrak{m}_i$.

Since the local maxima $\mathfrak{s}_1, \ldots, \mathfrak{s}_{n-1}$ are the boundaries of the potential wells Ω_i one can conclude from the construction (2.1.3) and (**S1**) that such a set $\mathcal{S}$ automatically does not contain these local maxima. Also note that (2.1.4) guarantees that (**S1**) and (**S2**) do not contradict each other.
The figure 2.1 illustrates this construction.
Additionally, for $1 \le i \le n$ let $\mathcal{S}_i$ be the part of $\mathcal{S}$ that belongs to the well Ω_i, i.e.

$$\mathcal{S}_i \coloneqq \mathcal{S} \cap \Omega_i.$$

Obviously the properties (**S1**) and (**S2**) also hold for each $\mathcal{S}_i$.
As an example for such a set $\mathcal{S}$ consider an equidistant decomposition of the interval $[-R,R]$ with step size $\delta = \frac{1}{M}$ for a large $M \in \mathbb{N}$, i.e. we set

$$J_1 \coloneqq \left[-R, -R + \frac{1}{M}\right),\ J_2 \coloneqq \left[-R + \frac{1}{M}, -R + \frac{2}{M}\right), \ldots, J_{2RM} \coloneqq \left[R - \frac{1}{M}, R\right].$$

Assume additionally that (2.1.4) holds for these intervals.
To get the intervals I_k one simply put $I_k = J_k$ except for those k for which there is an

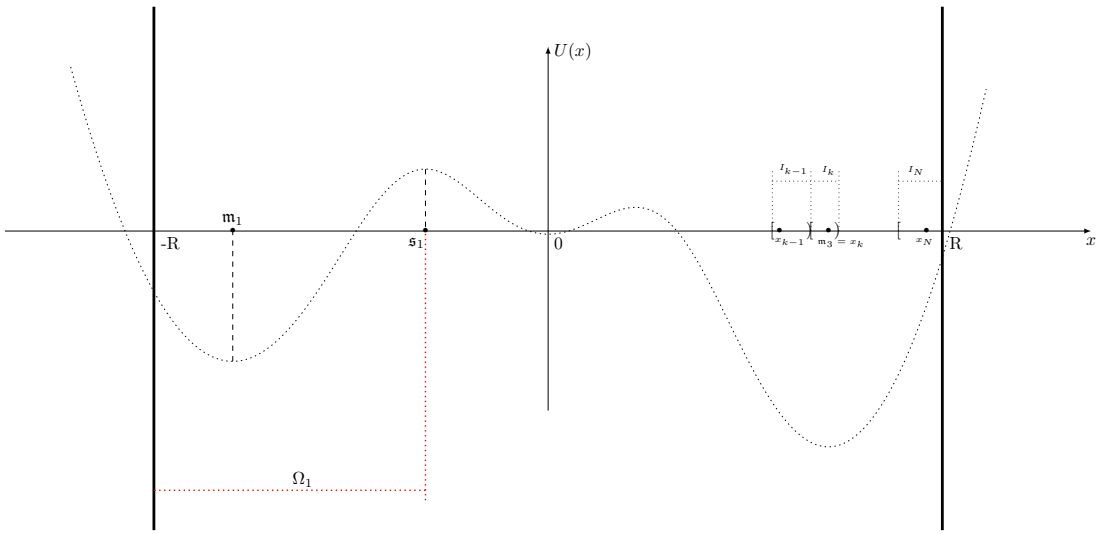


Figure 2.1: Construction of the state space according to (**S1**) and (**S2**).

overlap between J_k and two consecutive potential wells. Let l be the number of these overlaps, i.e.

$$l := \#\left\{1 \leq i \leq n-1 |\ \exists 1 \leq k \leq 2RM \text{ with } J_k \cap \Omega_i \neq \emptyset \text{ and } J_k \cap \Omega_{i+1} \neq \emptyset\right\}.$$

In this case just divide J_k into two intervals $I_k := J_k \cap \Omega_i$ and $I_{k+1} := J_k \cap \Omega_{i+1}$ and then continue by setting $I_{k+2} := J_{k+1}$. Altogether one obtains a sequence of intervals $I_1, \ldots, I_N$ with $N = 2RM + l$. Based on this one can construct a set $\mathcal{S}$ with the aforementioned properties by choosing the center of these intervals except for those which contain the local minimum. In this case one has to choose this minimum according to (**S2**).

2.2 Definition of the Markov Chain and Transition Probabilities

Now assume that a set $\mathcal{S}$ with the properties (**S1**) and (**S2**) is given. Our aim is to define two homogeneous Markov chains $Z^0(\cdot) = (Z_k^0(\cdot))_{k\geq 0}$ and $Z^\varepsilon(\cdot) = (Z_k^\varepsilon(\cdot))_{k\geq 0}$ on this set. Here the dot in the parenthesis stands for the starting point of the chains and will be specified below. Also, $Z^0(\cdot)$ will always correspond to the case $\varepsilon = 0$ and therefore will be referred to as *deterministic motion* or *deterministic Markov chain*.

Since $X_0^\varepsilon = x$, the initial distribution ν for both chains will be the Dirac distribution in

the point x_k for which $x \in I_k$ holds, i.e.

$$\nu(x_k) = \begin{cases} 1, & x \text{ and } x_k \text{ are in the same interval } I_k, \\ 0, & \text{otherwise.} \end{cases}$$

Now let us make the following agreements to ensure a certain clarity for the rest of this thesis.

Agreements:

(A1) From now on we will always identify the starting point x of the process X^ε with the starting point x_k of the chains Z^0 and Z^ε.

(A2) We will avoid the parenthesis that indicate the dependence on the starting point if the context allows us to do so, i.e. in these cases we will write Z^0 instead of $Z^0(\cdot)$ and so forth.

(A3) We will avoid the subscripts for the states of the Markov chains, i.e. we will just use the letters x, y, z instead of x_k, y_k, z_k.

(A4) Most of the quantities that are going to be introduced in this thesis will depend on the discretization parameters h, δ and R. But since we consider them as fixed we will suppress them whenever possible.

In view of these agreements we should adapt the definition of the interval boundaries given in (2.1.2), that is for $y \in \mathcal{S}$ we write from now on

$$\begin{aligned} a_y &\coloneqq \inf\{x|\ x \text{ and } y \text{ are in the same interval } I_k\} \quad \text{and} \\ b_y &\coloneqq \sup\{x|\ x \text{ and } y \text{ are in the same interval } I_k\}. \end{aligned} \tag{2.2.1}$$

By taking a closer look to (2.1.1) we get the idea for the definitions of the respective transition matrices. Again, since we think of Z^0 as the chain that belongs to the case $\varepsilon = 0$ we define its transition matrix $\mathbf{P}^0 = (p^0_{x,y})_{x,y\in\mathcal{S}}$ by

$$p^0_{x,y} \coloneqq \mathbb{P}(Z^0_1 = y | Z^0_0 = x) \coloneqq \begin{cases} 1, & a_y \le x - hU'(x) \le b_y, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, for a given $x \in \mathcal{S}$ let $y^*(x)$ be the uniquely determined state such that

$$p^0_{x,y^*(x)} = 1.$$

Now let us consider the chain Z^ε. Put $\tilde{y} \coloneqq \min\{y|\ y \in \mathcal{S}\}$ and $\hat{y} \coloneqq \max\{y|\ y \in \mathcal{S}\}$. Then the transition matrix $\mathbf{P}^\varepsilon = (p^\varepsilon_{x,y})_{x,y\in\mathcal{S}}$ is defined by

$$p^\varepsilon_{x,y} \coloneqq \mathbb{P}(a_y \le x - hU'(x) + \varepsilon h^{\frac{1}{\alpha}} L_1 \le b_y), \quad y \neq \tilde{y}, \hat{y},$$

as well as

$$p^\varepsilon_{x,\tilde{y}} \coloneqq \mathbb{P}(a_{\tilde{y}} \le x - hU'(x) + \varepsilon h^{\frac{1}{\alpha}} L_1 \le b_{\tilde{y}}) + \mathbb{P}(x - hU'(x) + \varepsilon h^{\frac{1}{\alpha}} L_1 \le -R),$$

$$p^\varepsilon_{x,\hat{y}} \coloneqq \mathbb{P}(a_{\hat{y}} \le x - hU'(x_k) + \varepsilon h^{\frac{1}{\alpha}} L_1 \le b_{\hat{y}}) + \mathbb{P}(x - hU'(x) + \varepsilon h^{\frac{1}{\alpha}} L_1 \ge R).$$

Note that in this model the Markov chain Z^ε will be reflected to the states $\tilde{y}$ and $\hat{y}$, respectively, whenever it tries to escape into the "outer world" $[-R, R]^c$.

To avoid technical difficulties we need to impose some additional conditions on the discretization parameters h and δ:

(D1) $(p^0_{x,x})^{(l)} = 0$ for every $l \ge 1$ and for every $x \in \mathcal{S} \setminus \mathcal{M}$ where $(p^0_{x,x})^{(l)}$ is the entry of the l-th step transition matrix.

(D2) For a given δ, and hence a given state space $\mathcal{S}$ with **(S1)** and **(S2)**, choose h such that for all $x, y \in \mathcal{S}$

$$a_y - x + hU'(x) \neq 0 \quad \text{and} \quad b_y - x + hU'(x) \neq 0.$$

Note that in the case $x \in \mathcal{M}$, which implies $U'(x) = 0$, this condition is still guaranteed by (2.1.4).

(D1) means that the deterministic chain is neither allowed to be trapped in a fixed state x except this state is a local minimum of U nor to have infinite loops. Let us remark that this requirement implies that we are not allowed to fix the parameter δ and then to choose

h arbitrarily small. **(D2)** will be particularly useful in Chapter 3 and guarantees that the chain Z^ε always tends to follow the trajectory of Z^0. It also implies that for a given $x \in \mathcal{S}$ the state $y^*(x)$ is uniquely characterized as the solution to the system of inequalities

$$a_{y^*(x)} - x + hU'(x) < 0 \quad \text{and} \quad b_{y^*(x)} - x + hU'(x) > 0. \tag{2.2.2}$$

Now let us calculate the probabilities $p^\varepsilon_{x,y}$ in the limit $\varepsilon \to 0$. For this recall the definition of $c_1(\alpha)$ in Proposition 1.4.1.

Proposition 2.2.1. *Let $0 < \varepsilon < \varepsilon_0$. Then there exists a constant $C = C(h, \delta, R, \alpha) > 0$ independent of $x, y \in \mathcal{S}$ such that the following estimates hold.*

(i) For $x, y \in \mathcal{S}$ with $y \neq \tilde{y}, \hat{y}, y^*(x)$ we have

$$\left| p^\varepsilon_{x,y} - d_{x,y}\, \varepsilon^\alpha \right| \leq C\varepsilon^{2\alpha} \tag{2.2.3}$$

where

$$d_{x,y} = \left| \frac{1}{|a_y - x + hU'(x)|^\alpha} - \frac{1}{|b_y - x + hU'(x)|^\alpha} \right| \frac{c_1(\alpha)h}{\alpha}.$$

(ii) For $x \in \mathcal{S}$ and $y \in \{\tilde{y}, \hat{y}\}$ we have

$$\left| p^\varepsilon_{x,y} - e_{x,y}\, \varepsilon^\alpha \right| \leq C\varepsilon^{2\alpha}$$

where

$$e_{x,y} = \begin{cases} d_{x,y} + \frac{c_1(\alpha)h}{\alpha|-R-x+hU'(x)|^\alpha}, & y = \tilde{y}, \\ d_{x,y} + \frac{c_1(\alpha)h}{\alpha|R-x+hU'(x)|^\alpha}, & y = \hat{y}. \end{cases}$$

(iii) For $x \in \mathcal{S}$ we have

$$\left| p^\varepsilon_{x,y^*(x)} - \left(1 - f_{x,y^*(x)}\, \varepsilon^\alpha\right) \right| \leq C\varepsilon^{2\alpha} \tag{2.2.4}$$

where

$$f_{x,y^*(x)} = \left(\frac{1}{\left|a_{y^*(x)} - x + hU'(x)\right|^\alpha} + \frac{1}{\left|b_{y^*(x)} - x + hU'(x)\right|^\alpha} \right) \frac{c_1(\alpha)h}{\alpha}.$$

Proof. For the proof of (i) fix x, y as mentioned in the formulation above and let

$$A^{\varepsilon}_{x,y} = \frac{a_y - x + hU'(x)}{h^{\frac{1}{\alpha}}\varepsilon} \quad \text{and} \quad B^{\varepsilon}_{x,y} = \frac{b_y - x + hU'(x)}{h^{\frac{1}{\alpha}}\varepsilon}.$$

Then

$$p^{\varepsilon}_{x,y} = \mathbb{P}(A^{\varepsilon}_{x,h} \leq L_1 \leq B^{\varepsilon}_{x,y}).$$

Because of $y \neq y^*(x)$, **(D2)** and (2.2.2) we have either $A^{\varepsilon}_{x,y}, B^{\varepsilon}_{x,y} < 0$ or $A^{\varepsilon}_{x,y}, B^{\varepsilon}_{x,y} > 0$. For the latter we find with Corollary 1.4.2

$$\left| p^{\varepsilon}_{x,y} - d_{x,y}(\alpha)\varepsilon^{\alpha} \right| \leq D_{x,y}(\alpha)\varepsilon^{2\alpha}$$

where

$$D_{x,y}(\alpha) = C(\alpha)h^2 \left(\frac{1}{|a_y - x + hU'(x)|^{2\alpha}} + \frac{1}{|b_y - x + hU'(x)|^{2\alpha}} \right).$$

Here $C(\alpha)$ is the constant from Corollary 1.4.2. Since the differences $a_y - x$ and $b_y - x$ can be viewed as functions of the parameter δ one can estimate

$$D_{x,y}(\alpha) \leq C(\alpha)h^2 \tilde{D}(\delta, h) \tag{2.2.5}$$

which proves (2.2.3).

The remaining cases follow with similar calculations by using the symmetry of L_1 and once again (2.2.2) in case (iii).

□

2.3 Asymptotics of the Stationary Distribution

In this section we study the asymptotic behavior of the stationary distribution $\boldsymbol{\pi}^{\varepsilon} = (\pi^{\varepsilon}_x)_{x \in \mathcal{S}}$ of Z^{ε}. Define the matrix $\mathbf{Q} = (q_{i,j})_{i,j=1}^n$ by

$$q_{i,j} = \begin{cases} \frac{1}{2} \left| \frac{1}{|\mathfrak{s}_{j-1} - \mathfrak{m}_i|^{\alpha}} - \frac{1}{|\mathfrak{s}_j - \mathfrak{m}_i|^{\alpha}} \right|, & i \neq j, \\ -\frac{1}{2} \left(\frac{1}{|\mathfrak{s}_i - \mathfrak{m}_i|^{\alpha}} + \frac{1}{|\mathfrak{s}_i - \mathfrak{m}_i|^{\alpha}} \right), & i = j, \end{cases}$$

This matrix is an infinitesimal generator of a time-continuous Markov chain $Y = (Y_t)_{t \geq 0}$ on the state space $\mathcal{M}$, see [ImPa2]. Denote by $\boldsymbol{\pi}^{\mathbf{Q}} = (\pi^{\mathbf{Q}}_i)_{i=1}^n$ the stationary distribution

of this chain.

The aim is to prove the following proposition.

Proposition 2.3.1. *The following statements hold:*

(i) For every $x \in \mathcal{S} \setminus \mathcal{M}$ we have

$$\lim_{\varepsilon \to 0} \pi_x^\varepsilon = 0.$$

(ii) For every $1 \leq i \leq n$ we have

$$\lim_{\varepsilon \to 0} \pi_{\mathfrak{m}_i}^\varepsilon = \pi_i^{\boldsymbol{Q}}.$$

Before we start with the proof let us ensure the uniqueness of both stationary distributions.

Lemma 2.3.2. *There exist unique stationary distributions for the Markov chains Z^ε and Y, i.e. the following linear systems possess a unique solution.*

$$(\mathbf{P}^\varepsilon)^T \boldsymbol{\pi}^\varepsilon = \boldsymbol{\pi}^\varepsilon, \quad \sum_{x \in \mathcal{S}} \pi_x^\varepsilon = 1, \quad \pi_x^\varepsilon \geq 0,\ x \in \mathcal{S}.$$

and

$$\mathbf{Q}^T \boldsymbol{\pi}^{\mathbf{Q}} = 0, \quad \sum_{i=1}^n \pi_i^{\mathbf{Q}} = 1, \quad \pi_i^{\mathbf{Q}} \geq 0,\ 1 \leq i \leq n.$$

Proof. From Proposition 2.2.1 we see that $p_{x,y}^\varepsilon > 0$ for every $\varepsilon > 0$ and hence we can apply Proposition 1.3.9. Also, by definition we have $|q_{i,j}| > 0$ and Proposition 1.3.12 then yields the statement for the Markov chain Y.

□

Usually it is not possible to conclude that a solution, if it exists, of a perturbed linear system $\mathbf{A}\mathbf{x}^\varepsilon = \mathbf{r}^\varepsilon$, $||\mathbf{r}^\varepsilon|| \to 0$, converges to a solution of $\mathbf{A}\mathbf{x} = \mathbf{0}$ if $\mathbf{A}$ is not invertible. However, in our case this conclusion is allowed.

Lemma 2.3.3. *Let $(\mathbf{r}^\varepsilon = (r_i^\varepsilon)_{i=1}^n)_{\varepsilon > 0}$ be a family of vectors such that $\max_{1 \leq i \leq n} |r_i^\varepsilon| \to 0$ as $\varepsilon \to 0$. Consider the system*

$$\mathbf{Q}^T \mathbf{s}^\varepsilon = \mathbf{r}^\varepsilon$$

and assume that this system has a solution for every $\varepsilon > 0$ *and that this solution is normalized such that* $\sum_{i=1}^{n} s_i^\varepsilon = 1$ *. Then, for all* $1 \leq i \leq n$ *we have*

$$\lim_{\varepsilon \to 0} s_i^\varepsilon = \pi_i^{\mathbf{Q}}.$$

Proof. Let us consider $\mathbf{Q}^T$ as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$ and let $E_1 \coloneqq \ker(\mathbf{Q}^T)$ and $E_2 \coloneqq \text{image}(\mathbf{Q}^T)$. We decompose $\mathbb{R}^n$ in the direct sum of E_1 and E_2, $\mathbb{R}^n = E_1 \oplus E_2$. Then it is easy to see that the restriction $\mathbf{Q}_2^T \coloneqq \mathbf{Q}^T|_{E_2}$ of $\mathbf{Q}^T$ on E_2 is a bijective mapping from E_2 onto E_2. Indeed, since every $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{Q}^T\mathbf{x} = 0$ is an element of E_1 we can deduce $\ker(\mathbf{Q}_2^T) = \{0\}$ and hence $\mathbf{Q}_2^T$ is injective. We also know that for every $\mathbf{x} \in E_2$ we can find $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2, \mathbf{y}_1 \in E_1, \mathbf{y}_2 \in E_2$ such that $\mathbf{x} = \mathbf{Q}^T\mathbf{y} = \mathbf{Q}^T\mathbf{y}_2 = \mathbf{Q}_2^T\mathbf{y}_2$ which implies the surjectivity.
Now let $\mathbf{s}^\varepsilon$ be the normalized solution to

$$\mathbf{Q}\mathbf{s}^\varepsilon = \mathbf{r}^\varepsilon.$$

We then can write $\mathbf{s}^\varepsilon = \mathbf{t}^\varepsilon + \mathbf{u}^\varepsilon$ for certain $\mathbf{t}^\varepsilon \in E_1,\ \mathbf{u}^\varepsilon \in E_2$. Hence

$$\mathbf{Q}^T\mathbf{u}^\varepsilon = \mathbf{r}^\varepsilon.$$

Since $\mathbf{Q}^T$ is invertible on E_2 and $(\mathbf{Q}^T)^{-1}|_{E_2} = (\mathbf{Q}_2^T)^{-1}|_{E_2}$ we can write $\mathbf{u}^\varepsilon = (\mathbf{Q}^\perp)^{-1}\mathbf{r}^\varepsilon$ and therefore $s_i^\varepsilon - t_i^\varepsilon = u_i^\varepsilon \to 0$ for every $1 \leq i \leq n$. Since $\mathbf{t}^\varepsilon \in E_1 = \{a\boldsymbol{\pi}^{\mathbf{Q}} |\ a \in \mathbb{R}\}$ we can conclude by the normalization condition on $\mathbf{s}^\varepsilon$ that

$$\lim_{\varepsilon \to 0} t_i^\varepsilon = \pi_i^{\mathbf{Q}}, \quad 1 \leq i \leq n,$$

which finishes the proof.

□

We are now in the position to prove Proposition 2.3.1.

Proof of Proposition 2.3.1. The first statement follows from the linear system of equations that determines the stationary distribution, i.e.

$$0 = \sum_{x \in \mathcal{S}} \pi_x^\varepsilon (\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})_{x,y}, \quad y \in \mathcal{S}. \tag{2.3.1}$$

(i) First consider states $y \in \mathcal{S} \setminus \mathcal{M}$ that cannot be reached by the deterministic motion, i.e. for which there is no state $\tilde{x}$ such that $y^*(\tilde{x}) = y$. That, for example, happens for the left and the right end of the state space, i.e. for $y = \tilde{y}$ and $y = \hat{y}$. In this case Proposition 2.2.1 (i) and (ii) yield

$$\lim_{\varepsilon \to 0} p^{\varepsilon}_{x,y} = 0 \quad \text{for all } x \in \mathcal{S},$$

in particular $p^{\varepsilon}_{y,y} \to 0$ as $\varepsilon \to 0$. Moreover, equation (2.3.1) can be read as

$$0 = \sum_{x \in \mathcal{S}} \pi^{\varepsilon}_x (\mathbf{P}^{\varepsilon} - \mathbf{I}_{|\mathcal{S}|})_{x,y} = \pi^{\varepsilon}_y (p^{\varepsilon}_{y,y} - 1) + \sum_{\substack{x \in \mathcal{S} \\ x \neq y}} \pi^{\varepsilon}_x p^{\varepsilon}_{x,y}.$$

But since π^{ε}_x is bounded for every x this forces π^{ε}_y to vanish in the limit $\varepsilon \to 0$.

(ii) Now let $y \in \mathcal{S} \setminus \mathcal{M}$ such that there exists $\tilde{x}$ with $y^*(\tilde{x}) = y$. Note that for such $\tilde{x}$ one always has $\tilde{x} \notin \mathcal{M}$. Collect all such states in the set $\mathcal{S}(y)$, i.e. $\mathcal{S}(y) \coloneqq \{\tilde{x} \in \mathcal{S} |\ y^*(\tilde{x}) = y\}$. Then, by Proposition 2.2.1 we have

$$\lim_{\varepsilon \to 0} p^{\varepsilon}_{\tilde{x},y} = 1 \quad \text{for all } \tilde{x} \in \mathcal{S}(y) \quad \text{and} \quad \lim_{\varepsilon \to 0} p^{\varepsilon}_{x,y} = 0 \quad \text{for all } x \notin \mathcal{S}(y).$$

Let us extend (2.3.1) a little bit further:

$$0 = \sum_{x \in \mathcal{S}} \pi^{\varepsilon}_x (\mathbf{P}^{\varepsilon} - \mathbf{I}_{|\mathcal{S}|})_{x,y} = \pi^{\varepsilon}_y (p^{\varepsilon}_{y,y} - 1) + \sum_{\tilde{x} \in \mathcal{S}(y)} \pi^{\varepsilon}_{\tilde{x}} p^{\varepsilon}_{\tilde{x},y} + \sum_{\substack{x \in \mathcal{S} \\ x \notin \mathcal{S}(y) \\ x \neq y}} \pi^{\varepsilon}_z p^{\varepsilon}_{z,y}.$$

By condition (**D1**) we have $y \notin \mathcal{S}(y)$ which implies $p^{\varepsilon}_{y,y} \to 0$ and hence

$$\lim_{\varepsilon \to 0} \left(\sum_{\tilde{x} \in \mathcal{S}(y)} \pi^{\varepsilon}_{\tilde{x}} p^{\varepsilon}_{\tilde{x},y} - \pi^{\varepsilon}_y \right) = 0. \tag{2.3.2}$$

Now we have to distinguish two cases. First let $\tilde{x} \in \mathcal{S}(y)$ such that there is no state $\hat{x}$ with $y^*(\hat{x}) = \tilde{x}$. Here we can apply the same argumentation as in (i) to derive

$$\lim_{\varepsilon \to 0} \pi^{\varepsilon}_{\tilde{x}} = 0.$$

In the second case we assume that there is a state $\hat{x}$ with $y^*(\hat{x}) = \tilde{x}$. Then by replacing y with $\tilde{x}$ we can use the same arguments that led to (2.3.2) to derive

$$\lim_{\varepsilon \to 0} \left(\sum_{\hat{x} \in \mathcal{S}(\tilde{x})} \pi^{\varepsilon}_{\hat{x}} p^{\varepsilon}_{\hat{x},\tilde{x}} - \pi^{\varepsilon}_{\tilde{x}} \right) = 0$$

Now we must perform the same case distinction again. This procedure stops after a finite number of steps since we will eventually end up in a state, say $\overline{x}$, that has no "predecessor" in the sense that there is no state x° such that $y^*(x^\circ) = \overline{x}$. Then we can successively conclude "forwards" again until we derive $\lim_{\varepsilon\to 0} \pi_y^\varepsilon = 0$.

To prove (ii) we consider a linear system which is equivalent to (2.3.1). For this we have to use some notations that we will introduce in more details in Chapter 4. Recall the definition of $c_1(\alpha)$ in Proposition 1.4.1 and set

$$c(\alpha, h) \coloneqq \frac{\alpha}{2c_1(\alpha)h} \quad \text{and} \quad a_{x,y}^\varepsilon \coloneqq \frac{c(\alpha,h)}{\varepsilon^\alpha}(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})_{x,y}.$$

We will show in Lemma 4.1.1 and Lemma 4.1.2 that

(a) For $x \notin \mathcal{M}$ we have $a_{x,x}^\varepsilon = -\frac{c(\alpha,h)}{\varepsilon^\alpha} + O(1)$, $a_{x,y^*(x)}^\varepsilon = \frac{c(\alpha,h)}{\varepsilon^\alpha} + O(1)$ and $a_{x,y}^\varepsilon = O(1)$ if $y \neq x, y^*(x)$.

(b) For $1 \leq i \leq n$ we have

$$\lim_{\varepsilon\to 0} \sum_{y\in\mathcal{S}_j} a_{\mathfrak{m}_i,y}^\varepsilon = q_{i,j}.$$

Now, consider the linear system

$$0 = \sum_{x\in\mathcal{S}} \pi_x^\varepsilon a_{x,y}^\varepsilon, \quad y \in \mathcal{S}. \tag{2.3.3}$$

Denote by $r_i \coloneqq |\mathcal{S}_i|$ the number of states in $\mathcal{S}_i$ and let us take a look at the first r_1 equations of the system (2.3.3). Fix a state $y \in \mathcal{S}_1 \setminus \{\mathfrak{m}_1\}$ and extend the equations that belong to the states y and $y^*(y)$, respectively:

$$0 = \pi_y^\varepsilon \left(-\frac{c(\alpha,h)}{\varepsilon^\alpha} + O(1)\right) + \sum_{i=1}^n \pi_{\mathfrak{m}_i}^\varepsilon a_{\mathfrak{m}_i,y}^\varepsilon + \sum_{\substack{x\in\mathcal{S}\setminus\mathcal{M}\\ x\neq y}} \pi_x^\varepsilon a_{x,y}^\varepsilon$$

$$0 = \pi_y^\varepsilon \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} + O(1)\right) + \pi_{y^*(y)}^\varepsilon \left(-\frac{c(\alpha,h)}{\varepsilon^\alpha} + O(1)\right) + \sum_{i=1}^n \pi_{\mathfrak{m}_i}^\varepsilon a_{\mathfrak{m}_i,y}^\varepsilon + \sum_{\substack{x\in\mathcal{S}\setminus\mathcal{M}\\ x\neq y,y^*(y)}} \pi_x^\varepsilon a_{x,y}^\varepsilon.$$

If we add these equations we see that the critical terms $-\pi_y^\varepsilon \frac{c(\alpha,h)}{\varepsilon^\alpha}$ and $\pi_y^\varepsilon \frac{c(\alpha,h)}{\varepsilon^\alpha}$ cancel out and hence after adding all r_1 equations up and using (a) and (i) we obtain an equation of the form

$$0 = \sum_{x \in \mathcal{S}_1 \setminus \mathfrak{m}_1} \pi_x^\varepsilon b_x^\varepsilon + \sum_{i=1}^{n} \pi_{\mathfrak{m}_i}^\varepsilon \left(\sum_{y \in \mathcal{S}_1} a_{\mathfrak{m}_i, y}^\varepsilon \right) + c^\varepsilon$$

where

$$b_x^\varepsilon = O(1) \text{ for } \in \mathcal{S}_1 \setminus \{\mathfrak{m}_1\} \quad \text{and} \quad c^\varepsilon = o(1).$$

This procedure can be done in each well $\mathcal{S}_j$. Using once more (i) and (b) we derive the following linear system in the limit $\varepsilon \to 0$:

$$\sum_{i=1}^{n} \pi_{\mathfrak{m}_i}^\varepsilon q_{i,j} = d_j^\varepsilon, \quad 1 \le j \le n,$$

where $d_j^\varepsilon = o(1)$. Finally, applying Lemma 2.3.3 finishes the proof.

□

3 Metastable Behavior of Z^ε

The aim of this chapter is to show that the Markov chain Z^ε defined in the previous chapter possesses a similar metastable behavior as the solution X^ε of equation (2.0.1). More precisely, consider the Q-matrix $\mathbf{Q} = (q_{i,j})_{i,j=1}^n$ given by

$$q_{i,j} = \begin{cases} \frac{1}{2}\left|\frac{1}{|\mathfrak{s}_{j-1}-\mathfrak{m}_i|^\alpha} - \frac{1}{|\mathfrak{s}_j-\mathfrak{m}_i|^\alpha}\right| & , \quad i \neq j, \\ -\frac{1}{2}\left(\frac{1}{|\mathfrak{s}_i-\mathfrak{m}_i|^\alpha} + \frac{1}{|\mathfrak{s}_i-\mathfrak{m}_i|^\alpha}\right), & \quad i = j, \end{cases} \tag{3.0.1}$$

and let $Y(\cdot) = (Y_t(\cdot))_{t\geq 0}$ be the continuous-time Markov chain with state space $\mathcal{M}$ generated by $\mathbf{Q}$. Note that the dot in the brackets again refers to the starting point of this Markov chain. Imkeller and Pavlyukevich proved in [ImPa2] the following theorem:

Theorem (Imkeller, Pavlyukevich '08).

If $x \in (\mathfrak{s}_{i-1}, \mathfrak{s}_i)$ for some $1 \leq i \leq n$, then

$$X^\varepsilon_{\frac{\alpha t}{2\varepsilon^\alpha}}(x) \to Y_t(\mathfrak{m}_i) \quad as \quad \varepsilon \to 0$$

in the sense of finite dimensional distributions.

Let us note that the authors in [ImPa2] atctually stated this theorem in a more general context where they allowed the Lévy measure of L to have a regular varying tail which includes α-stable processes.

The proof turned out to be relatively technical. In particular the dynamic in the neighborhood of an unstable saddle point had to be taken under special consideration. In our case we do not have to deal with this issue since we excluded the local maxima from the state space $\mathcal{S}$. Nonetheless, the proof of the following theorem remains rather technical and basically consists of the adaption of the ideas presented in [ImPa2] to our discrete setting.

Set

$$c(\alpha, h) \coloneqq \frac{\alpha}{2c_1(\alpha)h} \tag{3.0.2}$$

where $c_1(\alpha)$ is the constant given in Proposition 1.4.1.

Theorem 3.0.1. *Let $Z^\varepsilon(\cdot)$ be the discrete-time Markov chain defined in Chapter 2 and let $Y(\cdot)$ be the continuous-time Markov chain with generator $\mathbf{Q}$ defined in (3.0.1). Then, for every $t > 0$ and every $1 \le i, j \le n$ such that $i \neq j$ we have uniformly for $x \in \mathcal{S}_i$*

$$\lim_{\varepsilon \to 0} \mathbb{P}_x \left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t \rceil} = \mathfrak{m}_j \right) = \mathbb{P}_{\mathfrak{m}_i}(Y_t = \mathfrak{m}_j).$$

The main part of this chapter is devoted to the proof of this theorem. We start with the investigation of the first exit time of a single well. After that we consider the transition between the wells and finish the proof of Theorem 3.0.1. Note that we will also prove an additional result regarding the transition between two fixed wells which has no analog in [ImPa2].

3.1 Escape From a Single Well

In this section we focus on the dynamics in one well. For this recall in (2.1.1) the sequence $(\xi_k)_{k \ge 1}$ of independent and identical distributed random variables such that $\xi_1 \overset{d}{=} L_1$ and consider for fixed $1 \le i \le n$ a well $[c_i, d_i) \coloneqq \Omega_i$. For the sake of simplicity let us assume that 0 is the local minimum and that c_i, d_i are the local maxima of U.
Furthermore, let again $\mathcal{S}_i$ be the set of states from $\mathcal{S}$ that belong to Ω_i, i.e. $\mathcal{S}_i = \mathcal{S} \cap \Omega_i$. As we mentioned in Section 2.1 the properties (**S1**) and (**S2**) also hold for S_i. Together with the assumptions above this particularly implies that $0 \in \mathcal{S}_i$ and $c_i, d_i \notin \mathcal{S}_i$.
Also, recall the definitions (2.2.1) of the interval boundaries and of the index $y^*(x)$ for a given $x \in \mathcal{S}$. Define

$$D \coloneqq D(h, \delta) \coloneqq \min_{x \in \mathcal{S}} \min \left\{ \left| a_{y^*(x)} - (x - hU'(x)) \right| ; \left| b_{y^*(x)} - (x - hU'(x)) \right| \right\}. \tag{3.1.1}$$

This quantity gives the minimal distance which the deterministic motion has to the interval boundaries after one time step and therefore will be important when it comes to comparing the trajectories of Z^0 and Z^ε. Indeed, from now on we will refer to the event

$$\left\{\left|\varepsilon h^{\frac{1}{\alpha}}\xi_1\right| > D\right\}$$

as *big jump* since the occurrence of this event is a characterization for Z^0 and Z^ε to be in a different state after one time step, i.e.

$$Z_1^0(x) \neq Z_1^\varepsilon(x) \text{ for at least one } x \in \mathcal{S} \quad \Leftrightarrow \quad \left|\varepsilon h^{\frac{1}{\alpha}}\xi_1\right| > D.$$

Let us mention that the condition (**D2**) implies

$$0 < D < \delta.$$

Now we consider the times where big jumps occur and the time intervals between these jumps. Define recursively

$$\theta_0^\varepsilon := 0, \quad \theta_k^\varepsilon := \inf\left\{l > \theta_{k-1}^\varepsilon \mid \left|\varepsilon h^{\frac{1}{\alpha}}\xi_l\right| > D\right\}, \quad k \geq 1, \tag{3.1.2}$$

as well as

$$T_k^\varepsilon := \theta_k^\varepsilon - \theta_{k-1}^\varepsilon.$$

Moreover, let

$$p^\varepsilon := \mathbb{P}\left(\left|\varepsilon h^{\frac{1}{\alpha}}L_1\right| > D\right). \tag{3.1.3}$$

As a first result we prove the following easy lemma.

Lemma 3.1.1. *The following assertions hold:*

(i) For $k \geq 1$ one has $\mathbb{P}(T_1^\varepsilon = k) = p^\varepsilon(1-p^\varepsilon)^{k-1}$, i.e. T_1^ε is geometrically distributed with parameter p^ε.

(ii) $(T_k^\varepsilon)_{k\geq 1}$ is a sequence of independent and identically distributed random variables.

(iii) The random variables T_1^ε and $\xi_{T_1^\varepsilon}$ are independent.

Proof. The assertions (i) and (ii) are obvious. To prove (iii) we recall that the random variables $\xi_k,\ k \geq 1$, are independent and identically distributed. This yields for $v \in \mathbb{R}$

$$\begin{aligned}
\mathbb{E}\,\mathrm{e}^{iv\xi_{T_1^\varepsilon}} &= \sum_{k=1}^{\infty} \mathbb{E}\left[\mathrm{e}^{iv\xi_k}\,\mathbf{1}_{\{T_1^\varepsilon=k\}}\right] \\
&= \sum_{k=1}^{\infty} \mathbb{E}\left[\mathrm{e}^{iv\xi_k}\,\mathbf{1}_{\{|\varepsilon h^{1/\alpha}\xi_1|\leq D,\ldots,|\varepsilon h^{1/\alpha}\xi_{k-1}\leq D,|\varepsilon h^{1/\alpha}\xi_k|>D\}}\right] \\
&= \sum_{k=1}^{\infty} (1-p^\varepsilon)^{k-1}\mathbb{E}\left[\mathrm{e}^{iv\xi_k}\,\mathbf{1}_{\{|\varepsilon h^{1/\alpha}\xi_k|>D\}}\right] \\
&= \frac{\mathbb{E}\,\mathrm{e}^{iv\xi_1}\,\mathbf{1}_{\{|\varepsilon h^{1/\alpha}\xi_1|>D\}}}{p^\varepsilon}.
\end{aligned}$$

Furthermore, by (i) we know that $\mathbb{E}\,\mathrm{e}^{iuT_1^\varepsilon} = \frac{p^\varepsilon\,\mathrm{e}^{iu}}{1-(1-p^\varepsilon)e^{iu}},\ u \in \mathbb{R}$. Hence for the common characteristic function one finds with similar arguments

$$\begin{aligned}
\mathbb{E}\,\mathrm{e}^{iuT_1^\varepsilon+iv\xi_{T_1^\varepsilon}} &= \sum_{k=1}^{\infty} \mathbb{E}\left[\mathrm{e}^{iuk+iv\xi_k}\,\mathbf{1}_{\{T_1^\varepsilon=k\}}\right] \\
&= \sum_{k=1}^{\infty} \mathrm{e}^{iuk}(1-p^\varepsilon)^{k-1}\mathbb{E}\left[\mathrm{e}^{iv\xi_k}\,\mathbf{1}_{\{|\varepsilon h^{1/\alpha}\xi_k|>D\}}\right] \\
&= \frac{\mathbb{E}\left[\mathrm{e}^{iu\xi_1}\,\mathbf{1}_{\{|\varepsilon h^{1/\alpha}\xi_1|>D\}}\right]}{p^\varepsilon} p^\varepsilon\,\mathrm{e}^{iu}\sum_{k=0}^{\infty}\mathrm{e}^{iuk}(1-p^\varepsilon)^k \\
&= \mathbb{E}\,\mathrm{e}^{iv\xi_{T_1^\varepsilon}}\,\mathbb{E}\,\mathrm{e}^{iuT_1^\varepsilon}
\end{aligned}$$

which finishes the proof.

□

We are now interested in the behavior of the first exit time from the well $\mathcal{S}_i$. Define

$$\tau_i^\varepsilon := \inf\left\{k \geq 1 | Z_k^\varepsilon(\cdot) \notin \mathcal{S}_i\right\}. \tag{3.1.4}$$

Recall that in this section we assume that $\mathfrak{m}_i = 0$. Set

$$q_i^\varepsilon := \mathbb{P}(\varepsilon h^{1/\alpha}\xi_1 \notin \Omega_i). \tag{3.1.5}$$

and note that, as mentioned in agreement (**A4**), we will suppress the dependence on h of this quantity in the upcoming calculations.

With Corollary (1.4.2) we find

$$q_i^\varepsilon = \mathbb{P}(\varepsilon h^{1/\alpha}\xi_1 < c_i) + \mathbb{P}(\varepsilon h^{1/\alpha}\xi_1 > d_i) = \left(\frac{1}{|c_i|^\alpha} + \frac{1}{d_i^\alpha}\right)\frac{\varepsilon^\alpha}{c(\alpha,h)} + O\left(\varepsilon^{2\alpha}\right).$$

Now we can state the main result of this section which is an investigation of the limiting behavior of the standard Laplace transform of τ_i^ε. Denote by $\tau_i^\varepsilon(x)$ the first exit time conditioned on the starting point $x \in \mathcal{S}$. Furthermore, let $T^{\max}$ be the maximum time that the deterministic motion needs to reach a local minimum of U when starting from somewhere in $\mathcal{S}$, i.e.

$$T^{\max} := \max_{x\in\mathcal{S}} \min\left\{k \geq 1 |\ Z_k^0(x) \in \mathcal{M}\right\} \tag{3.1.6}$$

and note that this quantity obviously depends on the parameters h, δ and R.

Proposition 3.1.2. *Let $i = 1, \ldots, n$ and p^ε and q_i^ε be defined as above. Then the following assertions hold uniformly in $x \in \mathcal{S}_i$:*

(i) For every $u < 1$ we have

$$\lim_{\varepsilon\to 0} \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\right] = \frac{1}{1-u}$$

and hence $(q_i^\varepsilon\tau_i^\varepsilon(x))_{\varepsilon>0}$ converges in distribution to an Exp*(1)-distributed random variable.*

(ii) For every $p \geq 0$ we have

$$\lim_{\varepsilon\to 0} \mathbb{E}_x\left[q_i^\varepsilon\tau_i^\varepsilon\right]^p = \int_0^\infty y^p\,\mathrm{e}^{-y}\,dy.$$

Proof. Let us start with the proof of (i). Since the boundaries c_i, d_i of Ω_i do not belong to $\mathcal{S}_i$ and by definition of D one can conclude that the chain Z^ε can only escape from $\mathcal{S}_i$ if it makes a big jump. Therefore, with the notations introduced in (3.1.2), we can write for $u < 1$ and $\varepsilon < \varepsilon_0$, where u_0 and ε_0 are to be chosen later,

$$\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\right] = \sum_{k=1}^\infty \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\,\mathbf{1}_{\left\{\tau_i^\varepsilon=\theta_k^\varepsilon\right\}}\right]. \tag{3.1.7}$$

The main part of the proof consists of the estimate of these summands from above and below. We start with the latter. An iterative application of the strong Markov property as well as the use of Lemma 3.1.1 (ii) and (iii) yield

$$\begin{aligned}
\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\,\mathbf{1}_{\{\tau_i^\varepsilon=\theta_k^\varepsilon\}}\right] &= \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\theta_k^\varepsilon}\,\mathbf{1}_{\left\{Z_j^\varepsilon\in\mathcal{S}_i,1\le j\le\theta_k-1,Z_{\theta_k}^\varepsilon\notin\mathcal{S}_i\right\}}\right]\\
&= \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\theta_k^\varepsilon}\left(\prod_{j=1}^{k-1}\mathbf{1}_{\left\{Z_{\theta_j}^\varepsilon\in\mathcal{S}_i\right\}}\right)\mathbf{1}_{\left\{Z_{\theta_k}^\varepsilon\notin\mathcal{S}_i\right\}}\right]\\
&= \mathbb{E}_x\mathbb{E}\left[\mathrm{e}^{uq_i^\varepsilon\theta_k^\varepsilon}\left(\prod_{j=1}^{k-1}\mathbf{1}_{\left\{Z_{\theta_j}^\varepsilon\in\mathcal{S}_i\right\}}\right)\mathbf{1}_{\left\{Z_{\theta_k}^\varepsilon\notin\mathcal{S}_i\right\}}\,\middle|\,\mathcal{F}_{\theta_{k-1}^\varepsilon}\right]\\
&= \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\theta_{k-1}^\varepsilon}\left(\prod_{j=1}^{k-1}\mathbf{1}_{\left\{Z_{\theta_j}^\varepsilon\in\mathcal{S}_i\right\}}\right)\mathbb{E}\left(\mathrm{e}^{uq_i^\varepsilon(\theta_k^\varepsilon-\theta_{k-1}^\varepsilon)}\,\mathbf{1}_{\left\{Z_{\theta_k^\varepsilon}^\varepsilon\notin\mathcal{S}_i\right\}}\,\middle|\,Z_{\theta_{k-1}^\varepsilon}^\varepsilon\right)\right]\\
&= \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\theta_{k-1}^\varepsilon}\left(\prod_{j=1}^{k-1}\mathbf{1}_{\left\{Z_{\theta_j}^\varepsilon\in\mathcal{S}_i\right\}}\right)\mathbb{E}_{Z_{\theta_{k-1}^\varepsilon}^\varepsilon}\left(\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\notin\mathcal{S}_i\right\}}\right)\right]\\
&\ge \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\theta_{k-1}^\varepsilon}\left(\prod_{j=1}^{k-1}\mathbf{1}_{\left\{Z_{\theta_j}^\varepsilon\in\mathcal{S}_i\right\}}\right)\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left(\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\notin\mathcal{S}_i\right\}}\right)\right]\\
&= \mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\theta_{k-1}^\varepsilon}\left(\prod_{j=1}^{k-1}\mathbf{1}_{\left\{Z_{\theta_j}^\varepsilon\in\mathcal{S}_i\right\}}\right)\right]\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\notin\mathcal{S}_i\right\}}\right]\\
&\;\;\vdots\\
&\ge \left(\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\in\mathcal{S}_i\right\}}\right]\right)^{k-1}\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\notin\mathcal{S}_i\right\}}\right]. \qquad (3.1.8)
\end{aligned}$$

The indicator functions in (3.1.8) can be estimated as follows. Recall the definition of $T^{\max}$ in (3.1.6). Then the inclusion

$$\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[c_i,d_i]\right\}\cap\{T_1^\varepsilon\ge T^{\max}\}\subseteq\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}\cap\{T_1^\varepsilon\ge T^{\max}\}$$

holds and we find

$$\begin{aligned}
\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}} &\ge \mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}\cap\left\{T_1^\varepsilon\ge T^{\max}\right\}} \ge \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[c_i,d_i]\right\}\cap\left\{T_1^\varepsilon\ge T^{\max}\right\}}\\
&\ge \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[c_i,d_i]\right\}}-\mathbf{1}_{\left\{T_1^\varepsilon<T^{\max}\right\}}. \qquad (3.1.9)
\end{aligned}$$

With similar arguments we can estimate

$$\mathbf{1}_{\left\{Z^\varepsilon_{T^\varepsilon_1}(y)\notin\mathcal{S}_i\right\}} \geq \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\notin[c_i,d_i]\right\}\cap\left\{T^\varepsilon_1\geq T^{\max}\right\}} \geq \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\notin[c_i,d_i]\right\}}(1-\mathbf{1}_{\left\{T^\varepsilon_1<T^{\max}\right\}}). \tag{3.1.10}$$

Now, using (3.1.9), (3.1.10) and the independence of T^ε_1 and $\xi_{T^\varepsilon_1}$, we can continue estimating the expressions in the last line of (3.1.8):

$$\begin{aligned}\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\mathbf{1}_{\left\{Z^\varepsilon_{T^\varepsilon_1}\in\mathcal{S}_i\right\}}\right] &\geq \mathbb{E}\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\in[c_i,d_i]\right\}}\right]-\mathbb{E}\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\mathbf{1}_{\left\{T^\varepsilon_1<T^{\max}\right\}}\right]\\ &= \mathbb{E}\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\right]\mathbb{P}(\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\in[c_i,d_i])-\mathbb{E}\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\mathbf{1}_{\left\{T^\varepsilon_1<T^{\max}\right\}}\right].\end{aligned} \tag{3.1.11}$$

An easy calculation leads to

$$\mathbb{E}\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\right]=\sum_{k=1}^{\infty}\mathrm{e}^{uq_i^\varepsilon k}\,p^\varepsilon(1-p^\varepsilon)^{k-1}=\frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}} \tag{3.1.12}$$

while the last equation is justified as long as $u<\frac{-\ln(1-p^\varepsilon)}{q_i^\varepsilon}$. By (3.1.3) and (3.1.5) we have $\frac{p^\varepsilon}{q_i^\varepsilon}\geq 1$ for sufficiently small ε and δ and hence we can assume that (3.1.12) holds for $u<1$. Continuing in (3.1.11) yields

$$\begin{aligned}\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\mathbf{1}_{\left\{Z^\varepsilon_{T^\varepsilon_1}\in\mathcal{S}_i\right\}}\right] &\geq \frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\mathbb{P}(\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\in[c_i,d_i])-\\ &\qquad\frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\left(1-\left[(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}\right]^{T^{\max}-1}\right).\end{aligned}$$

Analogously,

$$\inf_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T^\varepsilon_1}\mathbf{1}_{\left\{Z^\varepsilon_{T^\varepsilon_1}\notin\mathcal{S}_i\right\}}\right]\geq\frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\mathbb{P}(\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\notin[c_i,d_i])\left[(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}\right]^{T^{\max}-1}.$$

Observe that $\mathbb{P}(\varepsilon h^{1/\alpha}\xi_{T^\varepsilon_1}\notin[c_i,d_i])=\frac{q_i^\varepsilon}{p^\varepsilon}$ and define

$$C_1^\varepsilon(u)\coloneqq 1-\left[(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}\right]^{T^{\max}-1}\quad\text{and}\quad C_{2,i}^\varepsilon(u)\coloneqq C_1^\varepsilon(u)\,\frac{p^\varepsilon}{q_i^\varepsilon}. \tag{3.1.13}$$

Inserting these estimates in (3.1.7) we find

$$\begin{aligned}
\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\right] &\geq \frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\frac{q_i^\varepsilon}{p^\varepsilon}(1-C_1^\varepsilon(u)) \\
&\quad\cdot\sum_{k=1}^{\infty}\left[\frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\left(1-\frac{q_i^\varepsilon}{p^\varepsilon}(1+C_{2,i}^\varepsilon(u))\right)\right]^{k-1} \\
&= \frac{\mathrm{e}^{uq_i^\varepsilon}\,q_i^\varepsilon(1-C_1^\varepsilon(u))}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\frac{1}{1-\frac{\mathrm{e}^{uq_i^\varepsilon}\,p^\varepsilon\left(1-\frac{q_i^\varepsilon}{p^\varepsilon}(1+C_{2,i}^\varepsilon(u))\right)}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}} \\
&= \frac{\mathrm{e}^{uq_i^\varepsilon}\,q_i^\varepsilon(1-C_1^\varepsilon(u))}{1-\mathrm{e}^{uq_i^\varepsilon}(1-q_i^\varepsilon(1+C_{2,i}^\varepsilon(u)))} \\
&= \frac{1}{\frac{1-\mathrm{e}^{uq_i^\varepsilon}}{\mathrm{e}^{uq_i^\varepsilon}\,q_i^\varepsilon(1-C_1^\varepsilon(u))}+\frac{1+C_{2,i}^\varepsilon(u)}{1-C_1^\varepsilon(u)}}.
\end{aligned} \tag{3.1.14}$$

Again, elementary manipulations together with $\frac{q_i^\varepsilon}{p^\varepsilon}\leq 1$ and $\lim_{\varepsilon\to 0}C_{2,i}^\varepsilon(u)=0$, $u\in\mathbb{R}$, show that the geometric series in the second step converges for every $u<1$ as long as ε and δ are sufficiently small.

Now let us estimate $\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\right]$ from above. Similarly to (3.1.8) we can find for this case

$$\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\,\mathbf{1}_{\{\tau_i^\varepsilon=\theta_k^\varepsilon\}}\right]\leq\left(\sup_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\in\mathcal{S}_i\right\}}\right]\right)^{k-1}\sup_{y\in\mathcal{S}_i}\mathbb{E}_y\left[\mathrm{e}^{uq_i^\varepsilon T_1^\varepsilon}\,\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon\notin\mathcal{S}_i\right\}}\right]. \tag{3.1.15}$$

The following inclusion holds for every $y\in\mathcal{S}_i$:

$$\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}\cap\{T_1^\varepsilon\geq T^{\max}\}\subseteq\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[c_i,d_i]\right\}\cap\{T_1^\varepsilon\geq T^{\max}\}.$$

Hence

$$\begin{aligned}
\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}} &= \mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}\cap\left\{T_1^\varepsilon\geq T^{\max}\right\}}+\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\in\mathcal{S}_i\right\}\cap\left\{T_1^\varepsilon<T^{\max}\right\}} \\
&\leq \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[c_i,d_i]\right\}}+\mathbf{1}_{\{T_1^\varepsilon<T^{\max}\}}.
\end{aligned}$$

Also, similar arguments yield

$$\mathbf{1}_{\left\{Z_{T_1^\varepsilon}^\varepsilon(y)\notin\mathcal{S}_i\right\}}\leq\mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\notin[c_i,d_i]\right\}}+\mathbf{1}_{\{T_1^\varepsilon<T^{\max}\}}.$$

Recalling the definitions of $C_1^\varepsilon(u)$ and $C_{2,i}^\varepsilon(u)$ in (3.1.13) and using once more the independence of T_1^ε and $\xi_{T_1^\varepsilon}$ as well as $u < 1$ we can continue in (3.1.15) and derive

$$\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\,\mathbf{1}_{\{\tau_i^\varepsilon=\theta_k^\varepsilon\}}\right] \le \left(\frac{p^\varepsilon\,\mathrm{e}^{uq_i^\varepsilon}}{1-(1-p^\varepsilon)\,\mathrm{e}^{uq_i^\varepsilon}}\right)^k \left(1-\frac{q_i^\varepsilon}{p^\varepsilon}(1-C_{2,i}^\varepsilon(u))\right)^{k-1}\left(\frac{q_i^\varepsilon}{p^\varepsilon}(1+C_{2,i}^\varepsilon(u))\right).$$

Inserting that into (3.1.7) yields

$$\mathbb{E}_x\left[\mathrm{e}^{uq_i^\varepsilon\tau_i^\varepsilon}\right] \le \frac{1}{\frac{1-\mathrm{e}^{uq_i^\varepsilon}}{\mathrm{e}^{uq_i^\varepsilon}\,q_i^\varepsilon(1+C_{2,i}^\varepsilon(u))}+\frac{1-C_{2,i}^\varepsilon(u)}{1+C_{2,i}^\varepsilon(u)}}. \tag{3.1.16}$$

Finally, passing to the limit $\varepsilon \to 0$ in (3.1.14) and (3.1.16) while using $\lim_{\varepsilon\to 0} C_1^\varepsilon(u) = \lim_{\varepsilon\to 0} C_{2,i}^\varepsilon(u) = 0$ proves (i).

For the proof of (ii) we observe that the convergence of the standard Laplace transform for at least one $u > 0$, which we just proved in (i), implies that all moments of $q_i^\varepsilon\tau_i^\varepsilon(x)$ are finite. Therefore, we can conclude$(q_i^\varepsilon\tau_i^\varepsilon(x))^p \Rightarrow V^p,\ \varepsilon \to 0$, where V is a Exp(1)-distributed random variable and that $((q_i^\varepsilon\tau_i^\varepsilon(x))^p)_{0<\varepsilon\le\varepsilon_0}$ is a family of uniformly integrable non-negative random variables. Applying Lemma 4.11 in [Ka] then yields the claim of (ii).

□

3.2 Transition Between the Wells

Let us consider the dynamics between the wells. Recall the decomposition of the state space

$$\mathcal{S} = \mathcal{S}_1 \cup \ldots \cup \mathcal{S}_n$$

where $\mathcal{S}_i = \mathcal{S} \cap \Omega_i$ and define for $1 \le i, j \le n$

$$q_{i,j}^\varepsilon := \mathbb{P}(\varepsilon h^{1/\alpha}\xi_1 \in [\mathfrak{s}_{j-1} - \mathfrak{m}_i, \mathfrak{s}_j - \mathfrak{m}_i]), \quad q_i^\varepsilon := \mathbb{P}(\varepsilon h^{1/\alpha}\xi_1 \notin [\mathfrak{s}_{i-1} - \mathfrak{m}_i, \mathfrak{s}_i - \mathfrak{m}_i]). \tag{3.2.1}$$

Note that the definition of q_i^ε is the same as in (3.1.5).

Clearly, if we assume that Z^ε starts in a local minimum $\mathfrak{m}_i$, then $q_{i,j}^\varepsilon$ is the probability of entering the well $\mathcal{S}_j$ and q_i^ε is the probability of escaping $\mathcal{S}_i$ at the first step.

Furthermore, recall the definition of $q_{i,j}$ in (3.0.1) and set $q_i := -q_{i,i}$.

Proposition 3.2.1. *Fix $1 \leq i, j \leq n$ and let τ_i^ε be the first exit time from the well $\mathcal{S}_i$ defined in (3.1.4). Then*

$$\lim_{\varepsilon \to 0} \mathbb{P}_x \left(Z^\varepsilon_{\tau_i^\varepsilon}(x) \in \mathcal{S}_j \right) = \frac{q_{i,j}}{q_i}$$

uniformly for $x \in \mathcal{S}_i$.

Proof. Since

$$\mathbb{P}_x \left(Z^\varepsilon_{\tau_i^\varepsilon}(x) \in \mathcal{S}_j \right) = \mathbb{E}_x \left[\mathrm{e}^{u q_i^\varepsilon \tau_i^\varepsilon} \mathbf{1}_{\left\{ Z^\varepsilon_{\tau_i^\varepsilon} \in \mathcal{S}_j \right\}} \right] \Bigg|_{u=0}$$

we can proceed very similarly to the proof of Proposition 3.1.2. That means we write for $x \in \mathcal{S}_i$ and $u < 1$

$$\mathbb{E}_x \left[\mathrm{e}^{u q_i^\varepsilon \tau_i^\varepsilon} \mathbf{1}_{\left\{ Z^\varepsilon_{\tau_i^\varepsilon} \in \mathcal{S}_j \right\}} \right] = \sum_{k=1}^{\infty} \mathbb{E}_x \left[\mathrm{e}^{u q_i^\varepsilon \theta_k} \mathbf{1}_{\left\{ Z^\varepsilon_{\tau_i^\varepsilon} \in \mathcal{S}_j \right\}} \mathbf{1}_{\{\tau_i^\varepsilon = \theta_k\}} \right]. \tag{3.2.2}$$

where $(\theta_k)_{k \geq 0}$ is the sequence of times of big jumps defined in (3.1.2).

Estimating these summands from below and above yields

$$\begin{aligned} \mathbb{E}_x \left[\mathrm{e}^{u q_i^\varepsilon \theta_k} \mathbf{1}_{\left\{ Z^\varepsilon_{\tau_i^\varepsilon} \in \mathcal{S}_j \right\}} \mathbf{1}_{\{\tau_i^\varepsilon = \theta_k\}} \right] \geq & \left(\inf_{y \in \mathcal{S}_i} \mathbb{E}_y \left[\mathrm{e}^{u q_i^\varepsilon T_1^\varepsilon} \mathbf{1}_{\left\{ Z^\varepsilon_{T_1^\varepsilon} \in \mathcal{S}_i \right\}} \right] \right)^{k-1} \\ & \cdot \inf_{y \in \mathcal{S}_i} \mathbb{E}_y \left[\mathrm{e}^{u q_i^\varepsilon T_1^\varepsilon} \mathbf{1}_{\left\{ Z^\varepsilon_{T_1^\varepsilon} \in \mathcal{S}_j \right\}} \right] \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}_x \left[\mathrm{e}^{u q_i^\varepsilon \theta_k} \mathbf{1}_{\left\{ Z^\varepsilon_{\tau_i^\varepsilon} \in \mathcal{S}_j \right\}} \mathbf{1}_{\{\tau_i^\varepsilon = \theta_k\}} \right] \leq & \left(\sup_{y \in \mathcal{S}_i} \mathbb{E}_y \left[\mathrm{e}^{u q_i^\varepsilon T_1^\varepsilon} \sup_{y \in \mathcal{S}_i} \mathbf{1}_{\left\{ Z^\varepsilon_{T_1^\varepsilon} \in \mathcal{S}_i \right\}} \right] \right)^{k-1} \\ & \cdot \sup_{y \in \mathcal{S}_i} \mathbb{E}_y \left[\mathrm{e}^{u q_i^\varepsilon T_1^\varepsilon} \mathbf{1}_{\left\{ Z^\varepsilon_{T_1^\varepsilon} \in \mathcal{S}_j \right\}} \right], \end{aligned}$$

respectively.

For these indicator functions we find

$$\begin{aligned} \mathbf{1}_{\left\{ Z^\varepsilon_{T_1^\varepsilon}(y) \in \mathcal{S}_i \right\}} &\geq \mathbf{1}_{\left\{ \varepsilon h^{1/\alpha} \xi_{T_1^\varepsilon} \in [\mathfrak{s}_{i-1} - \mathfrak{m}_i, \mathfrak{s}_i - \mathfrak{m}_i] \right\}} - \mathbf{1}_{\{T_1^\varepsilon < T^{\max}\}} \\ \mathbf{1}_{\left\{ Z^\varepsilon_{T_1^\varepsilon}(y) \in \mathcal{S}_j \right\}} &\geq \mathbf{1}_{\left\{ \varepsilon h^{1/\alpha} \xi_{T_1^\varepsilon} \in [\mathfrak{s}_{j-1} - \mathfrak{m}_i, \mathfrak{s}_j - \mathfrak{m}_i] \right\}} \left(1 - \mathbf{1}_{\{T_1^\varepsilon < T^{\max}\}} \right) \end{aligned}$$

as well as

$$\mathbf{1}_{\left\{Z^\varepsilon_{T_1^\varepsilon}(y)\in\mathcal{S}_i\right\}} \leq \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[\mathfrak{s}_{i-1}-\mathfrak{m}_i,\mathfrak{s}_i-\mathfrak{m}_i]\right\}} + \mathbf{1}_{\{T_1^\varepsilon<T^{\max}\}}$$
$$\mathbf{1}_{\left\{Z^\varepsilon_{T_1^\varepsilon}(y)\in\mathcal{S}_j\right\}} \leq \mathbf{1}_{\left\{\varepsilon h^{1/\alpha}\xi_{T_1^\varepsilon}\in[\mathfrak{s}_{j-1}-\mathfrak{m}_i,\mathfrak{s}_j-\mathfrak{m}_i]\right\}} + \mathbf{1}_{\{T_1^\varepsilon<T^{\max}\}}.$$

Now, using again the abbreviations $C_1^\varepsilon := \mathbb{P}(T_1^\varepsilon < T^{\max}) = 1-(1-p^\varepsilon)^{T^{\max}-1}$ and $C_{2,i}^\varepsilon := C_1^\varepsilon \frac{p^\varepsilon}{q_i^\varepsilon}$, we can insert these estimates into (3.2.2) for $u=0$ and derive

$$\mathbb{P}\left(Z^\varepsilon_{\tau_i^\varepsilon}(x)\in\mathcal{S}_j\right) \geq \frac{q_{i,j}^\varepsilon}{q_i^\varepsilon}\frac{1-C_1^\varepsilon}{1+C_{2,i}^\varepsilon}$$

and

$$\mathbb{P}\left(Z^\varepsilon_{\tau_i^\varepsilon}(x)\in\mathcal{S}_j\right) \leq \frac{q_{i,j}^\varepsilon}{q_i^\varepsilon}\frac{1+C_1^\varepsilon\frac{p^\varepsilon}{q_{i,j}^\varepsilon}}{1-C_{2,i}^\varepsilon}.$$

Finally, an easy calculation using Corollary 1.4.2 shows

$$\lim_{\varepsilon\to 0}\frac{q_{i,j}^\varepsilon}{q_i^\varepsilon} = \frac{q_{i,j}}{q_i}$$

which finishes the proof.

□

The next proposition shows that, in the limit $\varepsilon\to 0$, it makes not much of a difference to consider jumps between the wells or only the jumps between the local minima. Define

$$\eta_i^\varepsilon := \inf\left\{k\geq 1 |\ Z_k^\varepsilon\in\mathcal{M}\setminus\{\mathfrak{m}_i\}\right\}. \tag{3.2.3}$$

Proposition 3.2.2. *Let $1\leq i,j\leq n$ and $i\neq j$. Then*

$$\lim_{\varepsilon\to 0}\mathbb{P}_{\mathfrak{m}_i}\left(Z^\varepsilon_{\eta_i^\varepsilon}=\mathfrak{m}_j\right) = \frac{q_{i,j}}{q_i}.$$

Also, $(q_i^\varepsilon\eta_i^\varepsilon)_{\varepsilon>0}$ converges in distribution (with respect to $\mathbb{P}_{\mathfrak{m}_i}$) to an Exp(1)-distributed random variable.

Proof. Let again τ_i^ε be the first exit time of the well $\mathcal{S}_i$. Obviously one has

$$\mathbb{P}_{\mathfrak{m}_i}(\tau_i^\varepsilon\leq\eta_i^\varepsilon) = 1. \tag{3.2.4}$$

On the other hand, for the event $A_i^\varepsilon := \left\{\max_{1\le k\le T^{\max}} \left|\varepsilon h^{1/\alpha}\xi_{\tau_i^\varepsilon+k}\right| \le D\right\}$ we have $\lim_{\varepsilon\to 0}\mathbb{P}_{\mathfrak{m}_i}(A_i^\varepsilon) = 1$. That implies

$$\lim_{\varepsilon\to 0}\mathbb{P}_{\mathfrak{m}_i}\left(\eta_i^\varepsilon \le \tau_i^\varepsilon + T^{\max}\right) = 1. \tag{3.2.5}$$

Hence

$$\begin{aligned}\mathbb{P}_{\mathfrak{m}_i}\left(Z^\varepsilon_{\eta_i^\varepsilon} = \mathfrak{m}_j\right) &\ge \mathbb{P}_{\mathfrak{m}_i}\left(Z^\varepsilon_{\eta_i^\varepsilon} = \mathfrak{m}_j, Z^\varepsilon_{\tau_i^\varepsilon}\in\mathcal{S}_j, A_i^\varepsilon\right)\\ &= \mathbb{P}_{\mathfrak{m}_i}\left(Z^\varepsilon_{\tau_i^\varepsilon}\in\mathcal{S}_j, A_i^\varepsilon\right)\\ &\ge \mathbb{P}_{\mathfrak{m}_i}\left(Z^\varepsilon_{\tau_i^\varepsilon}\in\mathcal{S}_j\right) - \mathbb{P}_{\mathfrak{m}_i}\left((A_i^\varepsilon)^c\right)\end{aligned}$$

and the difference in the last line converges to $\frac{q_{i,j}}{q_i}$ as $\varepsilon\to 0$ by Proposition 3.2.1. Combining that with the equation

$$\sum_{\substack{j=1,\\ j\ne i}}^n \mathbb{P}_{\mathfrak{m}_i}\left(Z^\varepsilon_{\eta_i^\varepsilon}\in\mathcal{S}_j\right) = 1 = \sum_{\substack{j=1,\\ j\ne i}}^n \frac{q_{i,j}}{q_i}$$

finishes the proof of the first claim.

For the second assertion observe that similarly to (3.2.4) and (3.2.5) we have

$$\mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\tau_i^\varepsilon \le q_i^\varepsilon\eta_i^\varepsilon) = 1 \quad\text{and}\quad \lim_{\varepsilon\to 0}\mathbb{P}_{\mathfrak{m}_i}\left(q_i^\varepsilon\eta_i^\varepsilon \le q_i^\varepsilon\tau_i^\varepsilon + q_i^\varepsilon T^{\max}\right) = 1$$

and therefore, for $u\ge 0$,

$$\begin{aligned}|\mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\eta_i^\varepsilon > u) - \mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\tau_i^\varepsilon > u)| &\le |\mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\eta_i^\varepsilon > u) - \mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\tau_i^\varepsilon + q_i^\varepsilon T^{\max} > u)|\\ &\quad + |\mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\tau_i^\varepsilon + q_i^\varepsilon T^{\max} > u) - \mathbb{P}_{\mathfrak{m}_i}(q_i^\varepsilon\tau_i^\varepsilon > u)|\\ &\to 0, \quad \varepsilon\to 0.\end{aligned}$$

The claim then follows from Proposition 3.1.2.

□

We conclude this chapter by proving that the rescaled mean first hitting times of a fixed minimum $\mathfrak{m}_j\in\mathcal{M}$ for Z^ε started in $x\in\mathcal{S}_i, i\ne j$, converge to the mean first hitting time of $\mathfrak{m}_j$ of Y started in $\mathfrak{m}_i$.

For this define

$$m^{\mathbf{Q}}_{\mathfrak{m}_i,\mathfrak{m}_j} := \mathbb{E}_{\mathfrak{m}_i}\inf\{t\ge 0 |\, Y_t = \mathfrak{m}_j\}$$

and for $x, y \in \mathcal{S}$

$$m^\varepsilon_{x,y} = \mathbb{E}_x \inf\{k \geq 0 |\ Z^\varepsilon_k = y\}.$$

Recall the definition (3.0.2) of the scaling factor $c(\alpha, h)$.

Proposition 3.2.3. *Fix $1 \leq i, j \leq n$ such that $i \neq j$. Then*

$$\lim_{\varepsilon \to 0} \frac{\varepsilon^\alpha}{c(\alpha, h)} m^\varepsilon_{x,\mathfrak{m}_j} = m^{\mathbf{Q}}_{\mathfrak{m}_i,\mathfrak{m}_j}$$

uniformly for $x \in \mathcal{S}_i$.

Proof. Since the limits in Proposition 3.1.2 and Proposition 3.2.1 hold uniformly over the states of one well it suffices to show the claim for $x = \mathfrak{m}_i$.
We start by investigating the first hitting time $m^\varepsilon_{\mathfrak{m}_i,\mathcal{S}_j} = \mathbb{E}_{\mathfrak{m}_i} \inf\{k \geq 0 |\ Z^\varepsilon_k \in \mathcal{S}_j\}$ of the well $\mathcal{S}_j$. Let us introduce $(\Theta^{\mathbf{Q}}_l)_{l \geq 0}$ as the sequence of jump-times of Y and define by $(T^{\mathbf{Q}}_l)_{l \geq 0}$ the interval length between these jumps, i.e.

$$\Theta^{\mathbf{Q}}_0 := 0, \quad \Theta^{\mathbf{Q}}_l := \inf\left\{t > \Theta^{\mathbf{Q}}_{l-1} |\ Y_t \neq Y_{\Theta^{\mathbf{Q}}_{l-1}}\right\}, \quad l \geq 1,$$

and

$$T^{\mathbf{Q}}_l := \Theta^{\mathbf{Q}}_l - \Theta^{\mathbf{Q}}_{l-1}.$$

Analogously, define the respective quantities for Z^ε: For $x \in \mathcal{S}$ let $\mathcal{S}(x)$ denote the well that contains x, i.e. if $x \in \mathcal{S}_i$ then $\mathcal{S}(x) = \mathcal{S}_i$. Put

$$\Theta^\varepsilon_0 := 0, \quad \Theta^\varepsilon_l := \inf\left\{k > \Theta^\varepsilon_{l-1} |\ Z^\varepsilon_k \notin \mathcal{S}\left(Z^\varepsilon_{\Theta^\varepsilon_{k-1}}\right)\right\}, \quad T^\varepsilon_k := \Theta^\varepsilon_k - \Theta^\varepsilon_{k-1}.$$

First, consider the Markov chain Y. With the abbreviation $p^{\mathbf{Q}}_{i,j} = \mathbb{P}_{\mathfrak{m}_i}\left(Y_{T^{\mathbf{Q}}_1} = \mathfrak{m}_j\right)$ and

the strong Markov property we find

$$
\begin{aligned}
m^{\mathbf{Q}}_{\mathfrak{m}_i,\mathfrak{m}_j} &= \mathbb{E}_{\mathfrak{m}_i}(T^{\mathbf{Q}}_1)\, p^{\mathbf{Q}}_{i,j} \\
&\quad + \sum_{\substack{l=1\\ l\neq i,j}}^{n} \left(\mathbb{E}_{\mathfrak{m}_i}(T^{\mathbf{Q}}_1) + \mathbb{E}_{\mathfrak{m}_l}(T^{\mathbf{Q}}_2)\right) p^{\mathbf{Q}}_{i,l}\, p^{\mathbf{Q}}_{l,j} \\
&\quad + \sum_{\substack{l_1=1\\ l_1\neq i,j}}^{n} \sum_{\substack{l_2=1\\ l_2\neq l_1,j}}^{n} \left(\mathbb{E}_{\mathfrak{m}_i}(T^{\mathbf{Q}}_1) + \mathbb{E}_{\mathfrak{m}_{l_1}}(T^{\mathbf{Q}}_2) + \mathbb{E}_{\mathfrak{m}_{l_2}}(T^{\mathbf{Q}}_3)\right) p^{\mathbf{Q}}_{i,l_1}\, p^{\mathbf{Q}}_{l_1,l_2}\, p^{\mathbf{Q}}_{l_2,j} \\
&\quad + \dots \\
&= \mathbb{E}_{\mathfrak{m}_i}(T^{\mathbf{Q}}_1)\, p^{\mathbf{Q}}_{i,j} + \sum_{k=2}^{\infty} \left[\sum_{\substack{l_1=1\\ l_1\neq i,j}}^{n} \cdots \sum_{\substack{l_{k-1}=1\\ l\neq l_{k-2},j}}^{n} \left(\sum_{r=1}^{k} \mathbb{E}_{\mathfrak{m}_{l_{r-1}}}(T^{\mathbf{Q}}_r) \right) p^{\mathbf{Q}}_{i,l_1} \cdots p^{\mathbf{Q}}_{l_{k-1},j} \right]
\end{aligned}
$$

where we put $l_0 \coloneqq i$.

Since $\mathbb{E}_{\mathfrak{m}_i}(T^{\mathbf{Q}}_k) = O(1)$ for $1 \le i \le n$, $k \ge 1$ we can conclude that this series converges if and only if the series $\sum_{k=1}^{\infty} k f^{(k)}_{i,j}$ converges where

$$
f^{(k)}_{i,j} = \sum_{\substack{l_1=1\\ l_1\neq i,j}}^{n} \cdots \sum_{\substack{l_{k-1}=1\\ l_2\neq l_{k-2},j}}^{n} p^{\mathbf{Q}}_{i,l_1} \cdots p^{\mathbf{Q}}_{l_{k-1},j} = \mathbb{P}_{\mathfrak{m}_i}\left(Y_{\Theta^{\mathbf{Q}}_k} = \mathfrak{m}_j, Y_{\Theta^{\mathbf{Q}}_l} \neq \mathfrak{m}_j, 1 \le l \le k-1\right)
$$

and the latter is guaranteed by the result given in [Fr], Proposition 79, that states that there are constants $A > 0$ and $0 < r < 1$ such that

$$
f^{(k)}_{i,j} \le A r^k, \quad 1 \le i, j \le n. \tag{3.2.6}
$$

Similarly, for Z^{ε} we can write

$$
\begin{aligned}
&\frac{\varepsilon^{\alpha}}{c(\alpha,h)} m^{\varepsilon}_{\mathfrak{m}_i,\mathcal{S}_j} = \mathbb{E}_{\mathfrak{m}_i}\left(\frac{\varepsilon^{\alpha}}{c(\alpha,h)} T^{\varepsilon}_1\right) \mathbb{P}_{\mathfrak{m}_i}(Z^{\varepsilon}_{\Theta^{\varepsilon}_1} \in \mathcal{S}_j) \\
&+ \mathbb{E}_{\mathfrak{m}_i}\left[\sum_{\substack{l=1\\ l\neq i,j}}^{n} \left(\mathbb{E}_{\mathfrak{m}_i}\left(\frac{\varepsilon^{\alpha}}{c(\alpha,h)} T^{\varepsilon}_1\right) + \mathbb{E}_{\underbrace{Z^{\varepsilon}_{\Theta^{\varepsilon}_1}}_{\in \mathcal{S}_l}}\left(\frac{\varepsilon^{\alpha}}{c(\alpha,h)} T^{\varepsilon}_2\right) \right) \mathbb{P}_{\mathfrak{m}_i}(Z^{\varepsilon}_{\Theta^{\varepsilon}_1} \in \mathcal{S}_l) \mathbb{P}_{Z^{\varepsilon}_{\Theta^{\varepsilon}_1}}(Z^{\varepsilon}_{\Theta^{\varepsilon}_2} \in \mathcal{S}_j) \right] \\
&+ \cdots
\end{aligned}
$$

$$= \mathbb{E}_{\mathfrak{m}_i}\left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\Theta_1^\varepsilon\right)\mathbb{P}_{\mathfrak{m}_i}(Z^\varepsilon_{\Theta_1^\varepsilon}\in\mathcal{S}_j)$$
$$+\mathbb{E}_{\mathfrak{m}_i}\left[\sum_{k=2}^{\infty}\sum_{\substack{l_1=1\\ l_1\neq i,j}}^{n}\cdots\sum_{\substack{l_{k-1}=1\\ l\neq l_{k-2},j}}^{n}\left(\sum_{r=1}^{k}\mathbb{E}_{Z^\varepsilon_{\Theta^\varepsilon_{r-1}}}\left(\frac{\varepsilon^\alpha}{c(\alpha,h)}T_r^\varepsilon\right)\right)\prod_{r=1}^{k}\mathbb{P}_{Z^\varepsilon_{\Theta^\varepsilon_{r-1}}}(Z^\varepsilon_{\Theta^\varepsilon_r}\in\mathcal{S}_{l_r})\right] \quad (3.2.7)$$

where we put $l_k := j$. By Proposition 3.1.2 we know that $\mathbb{E}_{\mathfrak{m}_i}\left(\frac{\varepsilon^\alpha}{c(\alpha,h)}T_k^\varepsilon\right) = O(1)$. Furthermore, Proposition 3.2.1 implies the existence of a sequence $\kappa_{i,j}^\varepsilon$ such that $\lim_{\varepsilon\to 0}\kappa_{i,j}^\varepsilon = 0$ and $\mathbb{P}_x(Z^\varepsilon_{T_1^\varepsilon}\in\mathcal{S}_j) = p^{\mathbf{Q}}_{i,j}(1+\kappa^\varepsilon_{i,j})$ for $x\in\mathcal{S}_i$. We also know by Proposition 3.1.2 (ii) that, uniformly for $x\in\mathcal{S}_i$, $\lim_{\varepsilon\to 0}\mathbb{E}_x\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\Theta_1^\varepsilon\right] = \frac{1}{q_i} = \mathbb{E}_{\mathfrak{m}_i}\Theta_1^{\mathbf{Q}}$ and therefore there exists a sequence γ_i^ε with $\lim_{\varepsilon\to 0}\gamma_i^\varepsilon = 0$ such that $\mathbb{E}_x\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\Theta_1^\varepsilon\right] = \mathbb{E}_{\mathfrak{m}_i}\Theta_1^{\mathbf{Q}}(1+\gamma_i^\varepsilon)$. Then, the convergence of the series (3.2.7) follows from the convergence of

$$S^\varepsilon = \sum_{k=1}^{\infty} k f_{i,j}^{(k)}(1+\kappa^\varepsilon)^k$$

where $\kappa^\varepsilon = \max_{1\le i,j\le n}\kappa^\varepsilon_{i,j}$ which again follows by (3.2.6). Put $l_k = j$. Then we can continue in (3.2.7) to derive

$$\frac{\varepsilon^\alpha}{c(\alpha,h)}m^\varepsilon_{\mathfrak{m}_i,\mathcal{S}_j} = (1+\gamma_i^\varepsilon)\mathbb{E}_{\mathfrak{m}_i}(\Theta_1^{\mathbf{Q}})p^{\mathbf{Q}}_{i,j}(1+\kappa^\varepsilon_{i,j})$$
$$+\sum_{k=2}^{\infty}\sum_{\substack{l_1=1\\ l_1\neq i,j}}^{n}\cdots\sum_{\substack{l_{k-1}=1\\ l\neq l_{k-2},j}}^{n}\left(\sum_{r=1}^{k}\mathbb{E}_{\mathfrak{m}_{l_{r-1}}}(\Theta_1^{\mathbf{Q}})(1+\gamma^\varepsilon_{l_{r-1}})\right)\prod_{r=1}^{k}p^{\mathbf{Q}}_{l_{r-1},l_r}(1+\kappa^\varepsilon_{l_{r-1},l_r}). \quad (3.2.8)$$

Obviously, by definition of the numbers γ_i^ε and $\kappa^\varepsilon_{i,j}$ one has

$$\lim_{\varepsilon\to 0}\left(\sum_{r=1}^{k}\mathbb{E}_{\mathfrak{m}_{l_{r-1}}}(\Theta_1^{\mathbf{Q}})(1+\gamma^\varepsilon_{l_{r-1}})\right)\prod_{r=1}^{k}p^{\mathbf{Q}}_{l_{r-1},l_r}(1+\kappa^\varepsilon_{l_{s-1},l_s}) = \left(\sum_{r=1}^{k}\mathbb{E}_{\mathfrak{m}_{l_{r-1}}}(\Theta_1^{\mathbf{Q}})\right)\prod_{r=1}^{k}p^{\mathbf{Q}}_{l_{r-1},l_r}.$$

Now using the dominated convergence theorem in (3.2.8) yields the desired convergence $\lim_{\varepsilon\to 0}\frac{\varepsilon^\alpha}{c(\alpha,h)}m^\varepsilon_{\mathfrak{m}_i,\mathcal{S}_j} = m^{\mathbf{Q}}_{\mathfrak{m}_i,\mathfrak{m}_j}$.

To finish the proof we have to show that the rescaled mean exit time $\mathbb{E}_{\mathfrak{m}_i}\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon$ from the well $\mathcal{S}_i$ and the rescaled mean first hitting time $\mathbb{E}_{\mathfrak{m}_i}\frac{\varepsilon^\alpha}{c(\alpha,h)}\eta_i^\varepsilon$ of the set $\mathcal{M}\setminus\{\mathfrak{m}_i\}$ (see (3.2.3) for the definition) do not differ in the limit $\varepsilon\to 0$. Indeed, on one hand we have by definition

$$\mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon\right] \le \mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\eta_i^\varepsilon\right].$$

On the other hand recall the definition of the event A_i^ε from the proof of Proposition 3.2.2:

$$A_i^\varepsilon := \left\{ \max_{1\le k\le T^{\max}} \left| \varepsilon h^{1/\alpha} \xi_{\tau_i^\varepsilon + k} \right| \le D \right\}.$$

Then, for any given $\rho > 0$ we can find $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ one has $\mathbb{P}((A_i^\varepsilon)^c) \le \rho$ and, by Proposition 3.1.2, $\mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon\right] \le \frac{1}{q_i} + \rho$. It follows that

$$\begin{aligned}
\mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\eta_i^\varepsilon\right] &\le \mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon + \frac{\varepsilon^\alpha}{c(\alpha,h)}T^{\max}\right] \\
&\quad + \mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon + \frac{\varepsilon^\alpha}{c(\alpha,h)}2T^{\max}\right]\rho \\
&\quad + \dots \\
&\le \mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon\right] + \frac{\varepsilon^\alpha}{c(\alpha,h)}T^{\max} + \sum_{k=1}^{\infty}\left[\frac{1}{q_i} + \rho + \rho(k+1)T^{\max}\right]\rho^k
\end{aligned}$$

and therefore $\lim_{\varepsilon\to 0}\left|\mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\tau_i^\varepsilon\right] - \mathbb{E}_{\mathfrak{m}_i}\left[\frac{\varepsilon^\alpha}{c(\alpha,h)}\eta_i^\varepsilon\right]\right| = 0$. The claim of the proposition then follows by an application of Proposition 3.2.1 and Proposition 3.2.2 as well as similar arguments as we made above for $\mathbb{E}_{\mathfrak{m}_i,\mathcal{S}_j}$.

□

3.3 Proof of Theorem 3.0.1

First we show that on time scales which increase slower than $\varepsilon^{-\alpha}$ metastable behavior cannot be observed.

Proposition 3.3.1. *Fix $1 \le i \le n$ and let γ^ε be an increasing sequence such that*

$$\lim_{\varepsilon\to 0}\gamma^\varepsilon = \infty \quad \textit{and} \quad \lim_{\varepsilon\to 0}\varepsilon^\alpha\gamma^\varepsilon = 0.$$

Then, for all $x \in \mathcal{S}_i$ and $t > 0$, the sequence $\left(Z^\varepsilon_{\lceil\gamma^\varepsilon t\rceil}(x)\right)_{\varepsilon>0}$ converges in distribution to a degenerate random variable with the law $\delta_{\mathfrak{m}_i}$.

Proof. Fix $x \in \mathcal{S}_i$. We are going to show that for every $A > 0$ and sufficiently small $\rho > 0$

$$\lim_{\varepsilon\to 0}\mathbb{P}_x\left(\max_{T^{\max}\le t\le A\gamma^\varepsilon}\left|Z^\varepsilon_{\lceil t\rceil} - \mathfrak{m}_i\right| \le \rho\right) = 1$$

which then implies the claim of the proposition.

Indeed, using the Markov property we derive

$$\begin{aligned}\mathbb{P}_x\left(\max_{T^{\max}\le k\le A\gamma^\varepsilon}|Z_k^\varepsilon-\mathfrak{m}_i|>\rho\right) &\le \mathbb{P}_x\left(\max_{T^{\max}\le k\le A\gamma^\varepsilon}|Z_k^\varepsilon-\mathfrak{m}_i|>\rho,\ T_1^\varepsilon\ge T^{\max}\right)\\ &\quad+\mathbb{P}_x\left(T_1^\varepsilon<T^{\max}\right)\\ &\le \mathbb{P}_{\mathfrak{m}_i}\left(\max_{0\le k\le A\gamma^\varepsilon-T^{\max}}|Z_k^\varepsilon-\mathfrak{m}_i|>\rho\right)+\mathbb{P}_x\left(T_1^\varepsilon<T^{\max}\right)\\ &\le \mathbb{P}_{\mathfrak{m}_i}\left(\upsilon_{\rho,i}^\varepsilon\le A\gamma^\varepsilon-T^{\max}\right)+\mathbb{P}_x\left(T_1^\varepsilon<T^{\max}\right)\end{aligned}$$

where $\upsilon_{\rho,i}^\varepsilon := \min\{k\ge 0|\ |Z_k^\varepsilon-\mathfrak{m}_i|>\rho\}$. The second summand vanishes in the limit $\varepsilon\to 0$. Furthermore, with the definition $q_\rho^\varepsilon := \mathbb{P}(\left|\varepsilon h^{1/\alpha}\xi_1\right|\ge\rho)$ we can repeat the arguments of Proposition 3.1.2 to see that $(q_\rho^\varepsilon\upsilon_{\rho,i}^\varepsilon)_{\varepsilon>0}$ converges in distribution to an Exp(1)-distributed random variable and since $\lim_{\varepsilon\to 0}q_\rho^\varepsilon\gamma^\varepsilon=0$ this implies that the first summand also vanishes.

□

The next lemma states that Z^ε will end up in the set $\mathcal{M}$ with high probability if we scale the time axis with the factor $\frac{c(\alpha,h)}{\varepsilon^\alpha}$.

Lemma 3.3.2. *Fix $t>0$. Then we have uniformly for $x\in\mathcal{S}$*

$$\lim_{\varepsilon\to 0}\mathbb{P}_x\left(Z^\varepsilon_{\lceil\frac{c(\alpha,h)}{\varepsilon^\alpha}t\rceil}\in\mathcal{M}\right)=1.$$

Proof. Recall from (3.1.2) the definition of the first time a big jump occurs:

$$T_1^\varepsilon=\min\left\{k\ge 1|\ \left|\varepsilon h^{1/\alpha}\xi_k\right|>D\right\}.$$

where D is defined in (3.1.1). Then, by definition of $T^{\max}$, we find for all $x\in\mathcal{S}$

$$\mathbb{P}_x\left(Z^\varepsilon_{T^{\max}}\notin\mathcal{M}\right)\le\mathbb{P}(T_1^\varepsilon<T^{\max})$$

and thus

$$\lim_{\varepsilon\to 0}\mathbb{P}_x\left(Z^\varepsilon_{T^{\max}}\notin\mathcal{M}\right)=0.$$

Then, by applying the strong Markov property and the dominated convergence theorem we derive

$$\lim_{\varepsilon\to 0}\mathbb{P}_x\left(Z^\varepsilon_{\lceil\frac{c(\alpha,h)}{\varepsilon^\alpha}t\rceil}\notin\mathcal{M}\right)=\mathbb{E}_x\left[\lim_{\varepsilon\to 0}\mathbb{P}_{Z^\varepsilon_{\lceil\frac{c(\alpha,h)}{\varepsilon^\alpha}t\rceil-T^{\max}}}(Z^\varepsilon_{T^{\max}}\notin\mathcal{M})\right]=0$$

which proves the desired result.

□

Finally, we are in the position to prove Theorem 3.0.1.

Proof of Theorem 3.0.1. First, we define recursively a sequence of stopping times $(\vartheta^\varepsilon_l)_{l\geq 0}$ and states $(\mathfrak{m}^\varepsilon_l)_{l\geq 0}$ by $\vartheta^\varepsilon_0:=0$, $\mathfrak{m}^\varepsilon_0:=\mathfrak{m}_i$ and for $l\geq 1$

$$\vartheta^\varepsilon_l:=\min\left\{k\geq\vartheta^\varepsilon_{l-1}|\, Z^\varepsilon_k\in\mathcal{M}\setminus\{\mathfrak{m}_{l-1}\}\right\}\quad\text{and}\quad\mathfrak{m}^\varepsilon_l:=\sum_{\mu=1}^n\mathfrak{m}_\mu\mathbf{1}_{\left\{Z^\varepsilon_{\vartheta^\varepsilon_l}=\mathfrak{m}_\mu\right\}}.$$

Furthermore, let us define a process $(\tilde{Z}^\varepsilon_t)_{t\geq 0}$ on the state space $\mathcal{M}$ by

$$\tilde{Z}^\varepsilon_t:=\sum_{l=0}^\infty\mathfrak{m}^\varepsilon_l\mathbf{1}_{\left\{\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_l\leq t<\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_{l+1}\right\}}.$$

Note that this process is not necessarily Markovian. By definition we have $Z^\varepsilon_{\vartheta^\varepsilon_l}=\tilde{Z}^\varepsilon_{\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_l}$ for every $l\geq 0$ and from the strong Markov property as well as Proposition 3.2.2 it follows that in the limit $\varepsilon\to 0$ we have

$$\begin{aligned}&\text{Law}\left(\frac{\varepsilon^\alpha}{c(\alpha,h)}(\vartheta^\varepsilon_{l+1}-\vartheta^\varepsilon_l)|\,\tilde{Z}^\varepsilon_{\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_l}=\mathfrak{m}_i\right)\Rightarrow\text{Exp}(q^0_i)\quad\text{and}\\&\mathbb{P}_x\left(\tilde{Z}^\varepsilon_{\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_{l+1}}=\mathfrak{m}_j|\tilde{Z}^\varepsilon_{\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_l}=\mathfrak{m}_i\right)\to\frac{q_{i,j}}{q_i}\end{aligned}\tag{3.3.1}$$

uniformly in $l\geq 0$.

The process (Y_t) is uniquely characterized by its sequence of jump times and the states after these jumps. So consider again

$$\Theta^{\mathbf{Q}}_0:=0,\quad\Theta^{\mathbf{Q}}_l:=\inf\left\{t>\Theta^{\mathbf{Q}}_{l-1}|\,Y_t\neq Y_{\Theta^{\mathbf{Q}}_{l-1}}\right\},\ l\geq 1,$$

and

$$\overline{Y}_l:=Y_{\Theta^{\mathbf{Q}}_l},\ l\geq 0.$$

Then the random variables $\Theta^{\mathbf{Q}}_{l+1} - \Theta^{\mathbf{Q}}_l$, $l \geq 0$, are conditionally independent and exponentially distributed, i.e we have for every $l \geq 0$ and $1 \leq i, j \leq n,\ i \neq j$

$$\text{Law}\left(\Theta^{\mathbf{Q}}_{l+1} - \Theta^{\mathbf{Q}}_l | \ \overline{Y}_l = \mathfrak{m}_i\right) = Exp(q_i) \quad \text{and} \quad \mathbb{P}\left(\overline{Y}_{l+1} = \mathfrak{m}_j | \ \overline{Y}_l = \mathfrak{m}_i\right) = \frac{q_{i,j}}{q_i}. \tag{3.3.2}$$

We find

$$\begin{aligned}\left|\mathbb{P}_x\left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} = \mathfrak{m}_j\right) - \mathbb{P}_{\mathfrak{m}_i}(Y_t = \mathfrak{m}_j)\right| \leq &\left|\mathbb{P}_x\left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} = \mathfrak{m}_j\right) - \mathbb{P}_x\left(\tilde{Z}^\varepsilon_t = \mathfrak{m}_j\right)\right| \\ &+ \left|\mathbb{P}_x\left(\tilde{Z}^\varepsilon_t = \mathfrak{m}_j\right) - \mathbb{P}_{\mathfrak{m}_i}(Y_t = \mathfrak{m}_j)\right|\end{aligned}$$

The second summand vanishes in the limit $\varepsilon \to 0$ since the process $\tilde{Z}^\varepsilon$ converges weakly to Y. Indeed, in this case weak convergence is equivalent to the weak convergence of the sequence $(\frac{\varepsilon^\alpha}{c(\alpha,h)}\vartheta^\varepsilon_l, \mathfrak{m}^\varepsilon_l)_{l\geq 0} \Rightarrow (\Theta^{\mathbf{Q}}_l, \overline{Y}_l)_{l\geq 0}$ which follows from (3.3.1) and (3.3.2).
Moreover, by definition we have

$$\mathbb{P}_x\left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} = \mathfrak{m}_j\right) = \mathbb{P}_x\left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} = \mathfrak{m}_j, \tilde{Z}^\varepsilon_t = \mathfrak{m}_j\right)$$

as well as

$$\mathbb{P}_x\left(\tilde{Z}^\varepsilon_t = \mathfrak{m}_j\right) = \mathbb{P}_x\left(\tilde{Z}^\varepsilon_t = \mathfrak{m}_j, Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} = \mathfrak{m}_j\right) + \mathbb{P}_x\left(\tilde{Z}^\varepsilon_t = \mathfrak{m}_j, Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} \notin \mathcal{M}\right).$$

Then

$$\left|\mathbb{P}_x\left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} = \mathfrak{m}_j\right) - \mathbb{P}_{\mathfrak{m}_i}(Y_t = \mathfrak{m}_j)\right| \leq \mathbb{P}_x\left(Z^\varepsilon_{\lceil \frac{c(\alpha,h)}{\varepsilon^\alpha} t\rceil} \notin \mathcal{M}\right).$$

Finally, applying Lemma 3.3.2 yields the desired result.

□

4 Spectral Analysis for the Generator of Z^ε

The goal of this chapter is to present a detailed spectral analysis of the generator of the discrete Markov chain Z^ε and we start off by presenting the derivation of the matrix we are going to study. As mentioned in Section 1.3 the connection between the transition semigroup $\mathbf{P}_t$ of a homogeneous Markov process and the infinitesimal generator $\mathbf{Q}$ is given by the relation

$$\mathbf{P}_t = \mathrm{e}^{t\mathbf{Q}}\,.$$

In the case of a discrete-time Markov chain on a finite state space $\mathcal{S}$ the matrix $\mathbf{P} - \mathbf{I}_{|\mathcal{S}|}$, where $\mathbf{P}$ is the one step transition matrix, can be viewed as the analog for $\mathbf{Q}$, see Subsection 1.3.4. Now let again $\mathbf{P}^\varepsilon$ be the transition matrix of Z^ε. As explained in the introduction, the conjecture is that there are n eigenvalues of $\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|}$ which are of order $O(\varepsilon^\alpha)$ and that are separated from the remaining eigenvalues by a spectral gap. The main idea behind the proof of this claim is to use the results of Chapter 3 and to compare the spectrum of a suitable rescaled matrix with the spectrum $\sigma(\mathbf{Q})$ of the generator of the Markov chain Y. So let again $c(\alpha, h) \coloneqq \frac{\alpha}{2c_1(\alpha)h}$ and set

$$\mathbf{Q}^\varepsilon \coloneqq \frac{c(\alpha,h)}{\varepsilon^\alpha}\left(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|}\right). \tag{4.0.1}$$

This chapter is divided into three parts. First we will analyze the structure of the matrix

$$\mathbf{A}^\varepsilon(\lambda) \coloneqq \mathbf{Q}^\varepsilon - \lambda\mathbf{I}_{|\mathcal{S}|}, \quad \lambda \in \mathbb{C}, \tag{4.0.2}$$

which obviously plays the essential role in the investigation of the eigenvalues of $\mathbf{Q}^\varepsilon$ to which the second part is devoted. The last part then deals with the eigenvectors of the low lying eigenvalues.

4.1 The Structure of $\mathbf{A}^{\varepsilon}(\lambda)$

The first observation is that $\mathbf{A}^{\varepsilon}(\lambda)$ possesses a block structure. Indeed, since the decomposition $\mathcal{S} = \mathcal{S}_1 \sqcup \ldots \sqcup \mathcal{S}_n$ of the state space is disjoint one has

$$\mathbf{A}^{\varepsilon}(\lambda) = \left(\begin{array}{c|c|c} \mathbf{A}^{\varepsilon}_{\mathcal{S}_1,\mathcal{S}_1}(\lambda) & \cdots & \mathbf{A}^{\varepsilon}_{\mathcal{S}_1,\mathcal{S}_n} \\ \hline \vdots & & \vdots \\ \hline \mathbf{A}^{\varepsilon}_{\mathcal{S}_n,\mathcal{S}_1} & \cdots & \mathbf{A}^{\varepsilon}_{\mathcal{S}_n,\mathcal{S}_n}(\lambda) \end{array}\right).$$

Obviously only the blocks that include the main diagonal depend on λ. Since $\mathbf{A}^{\varepsilon}(\lambda) = \frac{c(\alpha,h)}{\varepsilon^{\alpha}}\left(\mathbf{P}^{\varepsilon} - \mathbf{I}_{|\mathcal{S}|}\right) - \lambda\mathbf{I}_{|\mathcal{S}|}$ one can see that an essential role is played by the transition matrix $\mathbf{P}^{\varepsilon}$ for which we can observe the same block structure. These blocks, let us call them $\mathbf{P}^{\varepsilon}_{\mathcal{S}_i,\mathcal{S}_j}$, determine the transition between the wells $\mathcal{S}_i$ and $\mathcal{S}_j$ or, in the case $i = j$, the transition within one well. With the help of Proposition 2.2.1 we can make the following statements about the entries $a^{\varepsilon}_{x,y}(\lambda)$.

Lemma 4.1.1. *The entries $a^{\varepsilon}_{x,y}(\lambda)$ are of the following order in the limit $\varepsilon \to 0$:*

(i) For $x, y \in \mathcal{S}$ and $y \notin \{x, y^*(x)\}$: $a^{\varepsilon}_{x,y} = O(1)$.

(ii) For $x \in \mathcal{S} \setminus \mathcal{M}$ and $y = y^*(x)$: $a^{\varepsilon}_{x,y^*(x)}(\lambda) = O(\varepsilon^{-\alpha})$. More precisely,

$$a^{\varepsilon}_{x,y^*(x)}(\lambda) = \frac{c(\alpha,h)}{\varepsilon^{\alpha}} - \tilde{a}^{\varepsilon}_{x,y^*(x)} \quad \text{where} \quad \tilde{a}^{\varepsilon}_{x,y^*(x)x} = O(1).$$

(iii) For $x \in \mathcal{S} \setminus \mathcal{M}$: $a^{\varepsilon}_{x,x}(\lambda) = O(\varepsilon^{-\alpha})$. More precisely,

$$a^{\varepsilon}_{x,x}(\lambda) = -\frac{c(\alpha,h)}{\varepsilon^{\alpha}} - \lambda + \tilde{a}^{\varepsilon}_{x,x} \quad \text{where} \quad \tilde{a}^{\varepsilon}_{x,x} = O(1).$$

(iv) For $x \in \mathcal{M}$: $a^{\varepsilon}_{x,x} = O(1)$. More precisely,

$$a^{\varepsilon}_{x,x} = -\tilde{a}^{\varepsilon}_{x,x} - \lambda \quad \text{where} \quad \tilde{a}^{\varepsilon}_{x,x} = O(1).$$

Proof. By Proposition 2.2.1 the entries of the transition matrix $p^{\varepsilon}_{x,y}$ are of order $O(\varepsilon^{\alpha})$ with the exception of the case $y = y^*(x)$. Hence, by multiplying with the scaling factor $\frac{c(\alpha,h)}{\varepsilon^{\alpha}}$ we have for $x, y \in \mathcal{S}$ such that $y \neq x$, $y^*(x)$

$$a^{\varepsilon}_{x,y} = O(1)$$

where, similar as in Proposition 2.2.1, the constant that appears in that Landau symbol can be chosen independent of x, y.
In the case $y = y^*(x)$ it holds that $p^\varepsilon_{x,y} = 1 - O(\varepsilon^\alpha)$. That yields the representation in (ii) and (iii). If $x \in \{\mathfrak{m}_1, \dots, \mathfrak{m}_n\}$ we have $y^*(x) = x$. Therefore, $(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})_{\mathfrak{m}_i,\mathfrak{m}_i} = O(\varepsilon^\alpha)$ which proves (iv).

□

For example let us take a look at the block $\mathbf{A}^\varepsilon_{\mathcal{S}_1,\mathcal{S}_1}(\lambda)$. If we number the states by $x_1, \dots, x_{N(1)}$, then this block has the form

$$\begin{pmatrix} -\frac{c(\alpha,h)}{\varepsilon^\alpha} - \lambda + \tilde{a}^\varepsilon_{x_1,x_1} & \cdots & \frac{c(\alpha,h)}{\varepsilon^\alpha} - \tilde{a}^\varepsilon_{x_1,y^*(x_1)} & \cdots & a^\varepsilon_{x_1,\mathfrak{m}_1} & \cdots & a^\varepsilon_{x_1,x_{N(1)}} \\ \vdots & & \vdots & & \cdots & & \vdots \\ a^\varepsilon_{\mathfrak{m}_1,x_1} & & a^\varepsilon_{\mathfrak{m}_1,y^*(x_1)} & & -\tilde{a}^\varepsilon_{\mathfrak{m}_1,\mathfrak{m}_1} - \lambda & & a^\varepsilon_{\mathfrak{m}_1,x_{N(1)}} \\ \vdots & & \vdots & & \cdots & & \vdots \\ a^\varepsilon_{x_{N(1)},x_1} & & a^\varepsilon_{x_{N(1)},y^*(x_1)} & & a^\varepsilon_{x_{N(1)},\mathfrak{m}_1} & \cdots & \frac{c(\alpha,h)}{\varepsilon^\alpha} - \lambda + \\ & & & & & & \tilde{a}^\varepsilon_{x_{N(1)},x_{N(1)}} \end{pmatrix}$$

Note that the numbers $\tilde{a}^\varepsilon_{x,y}$ that appear in the previous Lemma are all positive.

We conclude this section by showing that the summation over the entries of $\frac{c(\alpha,h)}{\varepsilon^\alpha}\mathbf{P}^\varepsilon$ indexed by a local minimum $\mathfrak{m}_i \in \mathcal{M}$ represents an approximation of the entries $q_{i,j}$ of $\mathbf{Q}$. Recall that

$$q_{i,j} = \frac{1}{2}\left|\frac{1}{|\mathfrak{s}_{j-1} - \mathfrak{m}_i|^\alpha} - \frac{1}{|\mathfrak{s}_j - \mathfrak{m}_i|^\alpha}\right|, \quad i \neq j.$$

Lemma 4.1.2. *Fix $1 \leq i, j \leq n$ such that $i \neq j$. Then there exists constants $\hat{C} = \hat{C}(h, \alpha, \delta, R)$ and $\varepsilon_0 > 0$ such that for every $0 < \varepsilon < \varepsilon_0$*

$$\left|q_{i,j} - \sum_{y \in \mathcal{S}_j} \frac{c(\alpha,h)}{\varepsilon^\alpha} p^\varepsilon_{\mathfrak{m}_i,y}\right| \leq \hat{C}\varepsilon^\alpha.$$

Proof. Recall that $\Omega_j = [\mathfrak{s}_{j-1}, \mathfrak{s}_j)$ with $\mathfrak{s}_0 = -\infty$ and $\mathfrak{s}_n = \infty$. An easy calculation shows that

$$q_{i,j} = \frac{\alpha}{2} \int_{\mathfrak{s}_{j-1}}^{\mathfrak{s}_j} \frac{1}{|y - \mathfrak{m}_i|^{\alpha+1}} dy.$$

First, consider the case $1 < j < n$ and let us write

$$\Omega_j = \bigsqcup_{x \in \mathcal{S}_j} [a_x, b_x)$$

where a_x, b_x are defined in (2.2.1). Then

$$q_{i,j} = \frac{\alpha}{2} \sum_{x \in \mathcal{S}_j} \int_{a_x}^{b_x} \frac{1}{|y - \mathfrak{m}_i|^{\alpha+1}} dy = \sum_{x \in \mathcal{S}_j} \frac{1}{2} \left| \frac{1}{|a_x - \mathfrak{m}_i|^{\alpha}} - \frac{1}{|b_x - \mathfrak{m}_i|^{\alpha}} \right|.$$

Moreover, by Proposition 2.2.1 (i) there exists a constant $\tilde{D} = \tilde{D}(\alpha, h, \delta)$ such that

$$\left| \frac{1}{2} \left| \frac{1}{|a_x - \mathfrak{m}_i|^{\alpha}} - \frac{1}{|b_x - \mathfrak{m}_i|^{\alpha}} \right| - \frac{c(\alpha, h)}{\varepsilon^{\alpha}} p^{\varepsilon}_{\mathfrak{m}_i, x} \right| \leq \tilde{D} \cdot \varepsilon^{\alpha}, \quad x \in \mathcal{S}_j,$$

and therefore

$$\left| q_{i,j} - \sum_{y \in \mathcal{S}_j} \frac{c(\alpha, h)}{\varepsilon^{\alpha}} p^{\varepsilon}_{\mathfrak{m}_i, x} \right| \leq |\mathcal{S}_j| \cdot \tilde{D} \cdot \varepsilon^{\alpha} \leq N \cdot \tilde{D} \cdot \varepsilon^{\alpha}.$$

A similar result can be derived for the cases $j = 1, n$ with the help of Proposition 2.2.1 (i) and (ii) although in this case the constant will additionally depend on R.

□

4.2 Eigenvalues

Recall that $|\mathcal{S}| = N$ and let $\lambda_1^{\mathbf{Q}}, \ldots, \lambda_n^{\mathbf{Q}}$ and $\lambda_1^{\varepsilon}, \ldots, \lambda_N^{\varepsilon}$ be the eigenvalues of $\mathbf{Q}$ and $\mathbf{Q}^{\varepsilon}$, respectively. We also use the notation $\sigma(\mathbf{B})$ to denote the spectrum of a matrix $\mathbf{B}$.
We start with a well known result.

Proposition 4.2.1. *All eigenvalues of* $\mathbf{Q}$ *and* $\mathbf{Q}^{\varepsilon}$ *have a non-positive real part.*

Proof. This result follows from Proposition A.2. Indeed, since $q_{i,i} < 0, 1 \leq i \leq n$, and $q^\varepsilon_{x,x} < 0, x \in \mathcal{S}$, as well as

$$\sum_{j \neq i} q_{i,j} = -q_{i,i} \quad \text{and} \quad \sum_{y \neq x} q^\varepsilon_{x,y} = -q^\varepsilon_{x,x}$$

all the Gershgorin disks are located in the negative half plane while the right end of each disk "touches" the origin which corresponds with the fact that $\lambda^\varepsilon_1 = \lambda^{\mathbf{Q}}_1 = 0$.

□

Let us now state to the main theorem of this chapter.

Theorem 4.2.2. *The spectrum $\sigma(\mathbf{Q}^\varepsilon)$ can be divided into two disjoint parts $\sigma_1(\mathbf{Q}^\varepsilon)$ and $\sigma_2(\mathbf{Q}^\varepsilon)$ for which the following assertions hold:*

(i) $\sigma_1(\mathbf{Q}^\varepsilon)$ contains precisely n eigenvalues $\lambda^\varepsilon_1, \ldots, \lambda^\varepsilon_n$ such that $\lambda^\varepsilon_1 = 0$ and

$$\lim_{\varepsilon \to 0} \lambda^\varepsilon_i = \lambda^{\mathbf{Q}}_i, \quad 2 \leq i \leq n.$$

(ii) There is a spectral gap that separates $\sigma_1(\mathbf{Q}^\varepsilon)$ and $\sigma_2(\mathbf{Q}^\varepsilon)$. More precisely, the following limit holds for the remaining eigenvalues $\lambda^\varepsilon_{n+1}, \ldots, \lambda^\varepsilon_N$:

$$\lim_{\varepsilon \to 0} |\lambda^\varepsilon_i| = \infty, \quad n+1 \leq i \leq N.$$

A simulation of such a Markov chain in *Mathematica 9* confirms the result of this theorem. As drift function we used $U'(x) = 10(x+1.8)(x+0.7)x(x-0.5)(x-1.6)/(1+x^4)$. A picture of an antiderivative of this function is shown in Section 2.1. The other parameters have the values $\alpha = 1.3$, $\varepsilon = 10^{-8}$, $h = \frac{1}{100}$, $\delta = \frac{1}{80}$ and $R = 2$ and $N = 320$. Figure 4.1 shows a plot of the spectrum. On the left hand side one can see the low lying eigenvalues and represents part (i) of Theorem 4.2.2 while the right picture shows part (ii). Note that we rescaled both axis such that the remaining eigenvalues $\lambda^\varepsilon_{n+1}, \ldots, \lambda^\varepsilon_N$ are of order $O(1)$.

The proof of this theorem mainly consists in the investigation of the characteristic polynomial $P^\varepsilon(\lambda) \coloneqq \det(\mathbf{A}^\varepsilon(\lambda))$ in the next proposition. We will show that, after a suitable

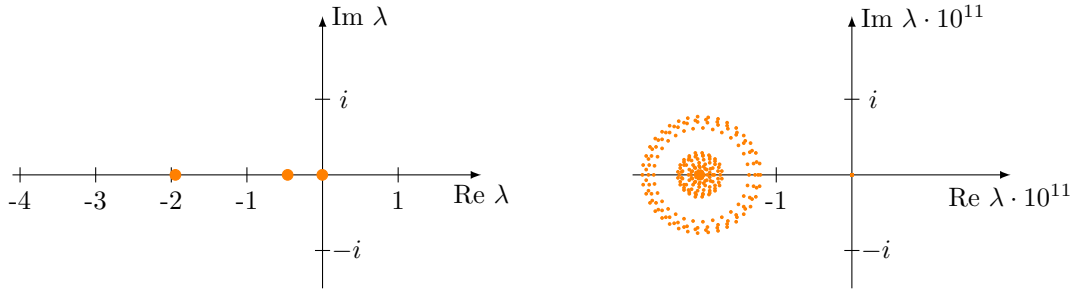


Figure 4.1: Left: Low lying eigenvalues; Right: The whole spectrum

scaling, it converges uniformly on a sufficiently large interval to the characteristic polynomial $P^{\mathbf{Q}}(\lambda) \coloneqq \det(\mathbf{A}^{\mathbf{Q}}(\lambda))$ where $\mathbf{A}^{\mathbf{Q}}(\lambda) \coloneqq \mathbf{Q} - \lambda \mathbf{I}_n$. Define $\hat{q} \coloneqq \max_{1\leq i\leq n} |q_{i,i}|$ and note that since $q_{i,i} = -\sum_{j\neq i} q_{i,j}$ and $q_{i,j} > 0, i \neq j$, one has $\hat{q} \geq |q_{i,j}|$ for all $1 \leq i, j \leq n$. Also, recall the constant $c(\alpha, h)$ defined in (3.0.2).

Proposition 4.2.3. *There exist constants $\varepsilon_0 > 0$, $C_1 = C_1(h, \delta, \alpha, R)$ and $C_2 = C_2(h, \delta, \alpha, R)$ such that for every $0 < \varepsilon < \varepsilon_0$ and $\lambda \in \mathbb{C}$*

$$\left| \left(-\frac{\varepsilon^\alpha}{c(\alpha, h)} \right)^{N-n} P^\varepsilon(\lambda) - P^{\mathbf{Q}}(\lambda) \right| \leq C_1 \max \left\{ (|\lambda| + C_2)^{n+1} ; \max_{1\leq k\leq n-1} (|\lambda| + \hat{q})^k \right\} \varepsilon^\alpha.$$

Before we start with the proof of this proposition let us show how it implies the statements of Theorem 4.2.2.

Proof of Theorem 4.2.2: Both statements are a consequence of Proposition 4.2.3 and the Hurwitz Theorem (see Proposition A.3). For the proof of (i) just choose a closed ball centered around 0 that includes all eigenvalues of $\mathbf{Q}$. Then Proposition 4.2.3 ensures uniform convergence inside this ball. But this proposition also provides us with a certain wiggle room when it comes to choosing the radius of this ball. Indeed, let $(z^\varepsilon)_{\varepsilon>0}$ be a sequence of complex numbers such that

$$\lim_{\varepsilon\to 0} |z^\varepsilon| = \infty \quad \text{and} \quad \lim_{\varepsilon\to 0} \varepsilon^\alpha |z^\varepsilon|^{n+1} = 0. \tag{4.2.1}$$

Put $r^\varepsilon := |z^\varepsilon|^{n+1}$. Then, Proposition 4.2.3 implies

$$\lim_{\varepsilon\to 0} \sup_{\lambda\in\overline{B}_{r^\varepsilon}(0)} \left| \left(-\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n} P^\varepsilon(\lambda) - P^{\mathbf{Q}}(\lambda) \right| = 0$$

and that proves the statement of (ii).

□

Let us remark that by (4.2.1) one can see that $\varepsilon^{-\frac{\alpha}{n+1}}$ is a lower bound for the increasing rate of the spectral gap.

The rest of this chapter is devoted to the proof of Proposition 4.2.3.
We start of with some preparations. Let $S_{|\mathcal{S}|} := \{\boldsymbol{\pi} = (\pi_x)_{x\in\mathcal{S}} |\ \boldsymbol{\pi} : \mathcal{S} \to \mathcal{S} \text{ bijective}\}$ be the set of all permutations of $\mathcal{S}$. Since $|\mathcal{S}| = N$ we will use the more common notation S_N from now on. The Leibniz formula for $P^\varepsilon(\lambda)$ then gives

$$P^\varepsilon(\lambda) = \sum_{\boldsymbol{\pi}\in S_N} \operatorname{sgn}(\boldsymbol{\pi}) \prod_{x\in\mathcal{S}} a^\varepsilon_{x,\pi_x}(\lambda). \tag{4.2.2}$$

Let us briefly explain the heuristic idea behind the following analysis of P^ε. Clearly, in view of Lemma 4.1.1 we have $|P^\varepsilon(\lambda)| \to \infty$ as $\varepsilon \to 0$ for any fixed $\lambda \in \mathbb{C}$. With that in mind let us ignore for a moment the values $\tilde{a}^\varepsilon_{x,y}$ since they are of order $O(1)$ with respect to $\varepsilon^{-\alpha}$. Then one can think of $P^\varepsilon(\lambda)$ as a polynomial of the variable $\varepsilon^{-\alpha}$ with coefficients depending on λ and terms of order $O(1)$. Also, keep in mind that these coefficients will depend on h, δ, R and α. This polynomial is of order $N - n$ (at most) since only the rows indexed by the minima $\mathfrak{m}_1, \ldots, \mathfrak{m}_n$ do not contain the term $\varepsilon^{-\alpha}$ and so we can write

$$P^\varepsilon(\lambda) = \sum_{j=0}^{N-n} c^\varepsilon_j(\lambda) \left(\varepsilon^{-\alpha}\right)^j.$$

Hence, by multiplying with $(\varepsilon^\alpha)^{N-n}$ we will "filter out" the coefficient $c^\varepsilon_{N-n}(\lambda)$, i.e. every other summand is going to vanish in the limit $\varepsilon \to 0$. But $c^\varepsilon_{N-n}(\lambda)$ contains expressions of the form $\sum_{y\in\mathcal{S}_j} \frac{c(\alpha,h)}{\varepsilon^\alpha} p^\varepsilon_{\mathfrak{m}_i,y}$, $1 \le i, j \le n$, which, as we saw in Lemma 4.1.2, are approximations of the entries $q_{i,j}$ of the generator matrix $\mathbf{Q}$. Finally, having that in mind we prove that $c^\varepsilon_{N-n}(\lambda)$ approximates $P^{\mathbf{Q}}$.

In view of the Leibniz formula (4.2.2) we should first turn to the permutations that contribute terms of order $O\left(\left(\varepsilon^{-\alpha}\right)^{N-n}\right)$ to the polynomial P^{ε}. This is what the next lemmas are about.

We see from Lemma 4.1.1 (ii) and (iii) that these permutations are precisely those which satisfy $\pi_x = x$ or $\pi_x = y^*(x)$ for $x \in \mathcal{S} \setminus \mathcal{M}$. So define

$$\Pi_0 \coloneqq \{\boldsymbol{\pi} = (\pi_x)_{x\in\mathcal{S}} \in S_N |\ \pi_x = x \text{ or } \pi_x = y^*(x) \text{ for all } x \in \mathcal{S} \setminus \mathcal{M}\}\,. \tag{4.2.3}$$

Lemma 4.2.4. *Fix $1 \le i \le n$. Then, for every $y \in \mathcal{S}_i$ there exists a permutation $\boldsymbol{\pi} \in \Pi_0$ such that*

$$\pi_{\mathfrak{m}_i} = y.$$

Proof. Fix $y \in \mathcal{S}_i$. If $y = \mathfrak{m}_i$, then we can choose $\boldsymbol{\pi} = \mathbf{id}$. Otherwise there exists a $k = k(y) \ge 1$ such that $((\mathbf{P}^0)^{(k)})_{y,\mathfrak{m}_i} = 1$, that is the deterministic motion, if started in y, reaches the minimum $\mathfrak{m}_i$ in k steps. Let us denote by $T : \mathcal{S} \to \mathcal{S}$ the mapping that belongs to the operation " $*$ ", i.e. $T(x) \coloneqq y^*(x)$. Then $((\mathbf{P}^0)^{(k)})_{y,\mathfrak{m}_i} = 1$ is equivalent to

$$(T^k)(y) \coloneqq (\underbrace{T \circ \ldots \circ T}_{k \text{ times}})(y) = \mathfrak{m}_i.$$

Now we can define a permutation $\hat{\boldsymbol{\pi}} : \mathcal{S} \to \mathcal{S}$ by

$$\hat{\pi}_y \coloneqq T(y),\ \hat{\pi}_{T(y)} = T^2(y), \ldots, \hat{\pi}_{T^{k-1}(y)} = T^k(y) = \mathfrak{m}_i,\ \hat{\pi}_{\mathfrak{m}_i} = y$$

as well as

$$\hat{\pi}_z = z \quad \text{if } z \in \mathcal{S} \setminus \left\{y, T(y), \ldots, T^k(y)\right\},$$

and from the construction it follows that $\hat{\boldsymbol{\pi}} \in \Pi_0$.

□

Let

$$\mathfrak{T}_n \coloneqq \{(y_1, \ldots, y_n) |\ y_i \in \mathcal{S}_i\}\,.$$

Lemma 4.2.5. *For every $(y_1, \dots, y_n) \in \mathfrak{T}_n$ there exists a permutation $\boldsymbol{\pi} \in \Pi_0$ such that*

$$\pi_{\mathfrak{m}_i} = y_i, \quad 1 \leq i \leq n.$$

Proof. This follows easily from the way we constructed the permutation in the proof of Lemma 4.2.4 since this construction only makes changes within $\mathcal{S}_i$ and these are disjoint so we can apply this principle for each of these wells simultaneously and glue the resulting permutations together.

□

Lemma 4.2.6. *For every $(y_1, \dots, y_n) \in \mathfrak{T}_n$ there exists a permutation $\boldsymbol{\pi} \in \Pi_0$ such that*

$$\pi_{\mathfrak{m}_i} = y_j, \quad 1 \leq i, j \leq n.$$

Proof. Fix $(y_1, \dots, y_n) \in \mathfrak{T}_n$ and an associated permutation $\boldsymbol{\pi}$ from Lemma 4.2.5 such that $\pi_{\mathfrak{m}_i} = y_i$. By definition of Π_0 the images of the states $\mathfrak{m}_1, \dots, \mathfrak{m}_n$ do not play a role when it comes to decide whether a permutation belongs to Π_0 or not and hence we can permute them among each other.

□

For the investigation of the determinant we need to introduce the following disjoint decomposition of the set S_N. For a given $0 \leq p \leq N - n$ let

$$\begin{aligned} \Pi_p \coloneqq \{\boldsymbol{\pi} \in S_N |\ & \text{There are precisely } p \text{ states } x_1, \dots, x_p \in \mathcal{S} \setminus \mathcal{M} \\ & \text{such that } \pi_{x_q} \notin \{x_q, y^*(x_q)\} \text{ for all } 1 \leq q \leq p\}. \end{aligned} \tag{4.2.4}$$

With this definition we can write

$$S_N = \bigsqcup_{p=0}^{N-n} \Pi_p.$$

Also note that the definition of Π_0 in (4.2.3) coincides with the one given in (4.2.4).
Now fix $0 \leq p \leq N - n$ and $\boldsymbol{\pi} \in \Pi_p$. We define

$$\begin{aligned} \mathcal{S}^1(\boldsymbol{\pi}) &\coloneqq \{x \in \mathcal{S} \setminus \mathcal{M} |\ \pi_x = x\}, \\ \mathcal{S}^2(\boldsymbol{\pi}) &\coloneqq \{x \in \mathcal{S} \setminus \mathcal{M} |\ \pi_x = y^*(x)\}, \\ \mathcal{S}^3(\boldsymbol{\pi}) &\coloneqq \mathcal{S} \setminus \left(\mathcal{S}^1(\boldsymbol{\pi}) \cup \mathcal{S}^2(\boldsymbol{\pi})\right). \end{aligned}$$

The following Lemma follows directly from the definition of Π_p.

Lemma 4.2.7. *Fix $0 \leq p \leq N-n$.*

(i) For every $\boldsymbol{\pi} \in \Pi_0$ we have $\mathcal{S}^3(\boldsymbol{\pi}) = \mathcal{M}$.

(ii) For every $\boldsymbol{\pi} \in \Pi_p$ we have $|\mathcal{S}^1(\boldsymbol{\pi}) \sqcup \mathcal{S}^2(\boldsymbol{\pi})| = N-n-p$.

Furthermore, let us introduce functions f_p that count the number of fixed points of $\boldsymbol{\pi} \in \Pi_p$ on the set $\mathcal{S} \setminus \mathcal{M}$, i.e.

$$f_p(\boldsymbol{\pi}) \coloneqq |\{x \in \mathcal{S} \setminus \mathcal{M} |\ \pi_x = x\}| = |\mathcal{S}^1(\boldsymbol{\pi})|.$$

Note that by definition of Π_p it follows that f_p has values in $\{0, \dots, N-n-p\}$.
We then have

$$\prod_{x \in \mathcal{S}^1(\boldsymbol{\pi})} \left(-\frac{c(\alpha,h)}{\varepsilon^\alpha} - \lambda + \tilde{a}^\varepsilon_{x,x} \right) = (-1)^{f_p(\boldsymbol{\pi})} \prod_{x \in \mathcal{S}^1(\boldsymbol{\pi})} \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} + (\lambda - \tilde{a}^\varepsilon_{x,x}) \right)$$

which allows us to write the characteristic polynomial P^ε as

$$\begin{aligned} P^\varepsilon(\lambda) = \sum_{p=0}^{N-n} \Bigg[\sum_{\boldsymbol{\pi} \in \Pi_p} \operatorname{sgn}(\boldsymbol{\pi}) (-1)^{f_p(\boldsymbol{\pi})} &\cdot \prod_{x \in \mathcal{S}^1(\boldsymbol{\pi})} \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} + (\lambda - \tilde{a}^\varepsilon_{x,x}) \right) \\ &\cdot \prod_{x \in \mathcal{S}^2(\boldsymbol{\pi})} \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} - \tilde{a}^\varepsilon_{x,y^*(x)} \right) \\ &\cdot \prod_{x \in \mathcal{S}^3(\boldsymbol{\pi})} a^\varepsilon_{x,\pi_x} \Bigg] \end{aligned} \tag{4.2.5}$$

Using Lemma 4.2.7 (ii) together with a simple straightforward calculation shows that there are numbers $\beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda)$ such that

$$\prod_{x \in \mathcal{S}^1(\boldsymbol{\pi})} \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} + (\lambda - \tilde{a}^\varepsilon_{x,x}) \right) \prod_{x \in \mathcal{S}^2(\boldsymbol{\pi})} \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} - \tilde{a}^\varepsilon_{x,x} \right) = \sum_{l=0}^{N-n-p} \beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda) \left(\frac{c(\alpha,h)}{\varepsilon^\alpha} \right)^l. \tag{4.2.6}$$

Note that the coefficients $\beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda)$ are of order $O(1)$ with respect to ε.
As explained before P^ε can roughly be viewed as a polynomial in the variable $\varepsilon^{-\alpha}$ and our aim is to isolate the coefficient that belongs to the highest power. So after scaling with

the factor $(\frac{\varepsilon^\alpha}{c(\alpha,h)})^{N-n}$ and taking Lemma 4.2.7 (i) into account we split the sum in (4.2.5) up and derive

$$\begin{aligned}\left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n} P^\varepsilon(\lambda) &= \sum_{\boldsymbol{\pi}\in\Pi_0} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} \prod_{i=1}^{n} a^\varepsilon_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) \\ &+ \sum_{\boldsymbol{\pi}\in\Pi_0} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} \left(\sum_{l=0}^{N-n-1} \beta^\varepsilon_{l,0,\boldsymbol{\pi}}(\lambda) \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l}\right) \prod_{i=1}^{n} a^\varepsilon_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) \\ &+ \sum_{p=1}^{N-n} \sum_{\boldsymbol{\pi}\in\Pi_p} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_p(\boldsymbol{\pi})} \left(\sum_{l=0}^{N-n-p} \beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda) \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l}\right) \prod_{x\in\mathcal{S}^3(\boldsymbol{\pi})} a^\varepsilon_{x,\pi_x}.\end{aligned} \tag{4.2.7}$$

Let us continue the investigation of the permutations.

Lemma 4.2.8. *Fix $\boldsymbol{\pi} \in \Pi_0$ such that $\pi_{\mathfrak{m}_i} \in \mathcal{S}_i$ for every $1 \le i \le n$. Then*

$$\operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} = (-1)^{N-n}.$$

Proof. Fix $1 \le i \le n$ and let $\pi_{\mathfrak{m}_i} = y_i \in \mathcal{S}_i$. From this condition and from the definition of Π_0 it follows that for every $x \in \mathcal{S}_i$ we have $\pi_x \in \mathcal{S}_i$. Hence, we can write $\boldsymbol{\pi}$ as a product of disjoint cycles,

$$\boldsymbol{\pi} = \xi_1 \cdot \ldots \cdot \xi_n,$$

where each cycle ξ_i acts only on $\mathcal{S}_i$. Define $k_i = k(y_i) \ge 0$ as the number of steps the deterministic motion needs to reach $\mathfrak{m}_i$ if started in y_i (see the proof of Lemma 4.2.4). Then we can write each cycle ξ_i as

$$\xi_i = \begin{cases} (y_i\, T(y_i) \cdots T^{k_i-1}(y_i)\, \mathfrak{m}_i), & k_i \ge 1, \\ (\mathfrak{m}_i), & k_i = 0. \end{cases}$$

(Again, for the notation see the proof of Lemma 4.2.4) The case $k_i = 0$ happens when $\boldsymbol{\pi}$ acts as identity in $\mathcal{S}_i$, i.e. $\pi_x = x$ for all $x \in \mathcal{S}_i$. Note that, against the usual convention, we do not omit cycles of length 1.

Let $|\xi_i|$ denote the length of the i-th cycle. Then $|\xi_i| = k_i + 1$ and

$$\operatorname{sgn}(\boldsymbol{\pi}) = (-1)^{|\xi_1|-1} \cdot \ldots \cdot (-1)^{|\xi_n|-1} = (-1)^{\sum_{i=1}^n k_i}.$$

On the other hand, from the definition of f_0 it follows that

$$f_0(\boldsymbol{\pi}) = \sum_{i=1}^{n} (|\mathcal{S}_i| - (k_i + 1)) = N - n - \sum_{i=1}^{n} k_i$$

which proves the claim.

□

We decompose the set Π_0. For this let S_n denote the set of all permutations $\boldsymbol{\sigma} = (\sigma_i)_{i=1}^n$: $\{1, \dots, n\} \to \{1, \dots, n\}$ and set

$$\Pi(\boldsymbol{\sigma}) \coloneqq \{\boldsymbol{\pi} \in \Pi_0 |\ \pi_{\mathfrak{m}_i} \in \mathcal{S}_{\sigma_i},\ 1 \leq i \leq n\}\,.$$

Then it is easy to see that

$$\Pi_0 = \bigsqcup_{\boldsymbol{\sigma} \in S_n} \Pi(\boldsymbol{\sigma}).$$

Lemma 4.2.9. *Let $\boldsymbol{\sigma} \in S_n$. Then, for every $\boldsymbol{\pi} \in \Pi(\boldsymbol{\sigma})$*

$$\operatorname{sgn}(\boldsymbol{\sigma})(-1)^{N-n} = \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})}.$$

Proof. Let $\boldsymbol{\sigma} = \zeta_1 \cdot \ldots \cdot \zeta_L$ be a representation of $\boldsymbol{\sigma}$ as a product of disjoint cycles where we again count cycles of length 1. In particular that implies $\sum_{l=1}^{L} |\zeta_l| = n$.
Moreover, from the definition of $\Pi(\sigma)$ it follows that there exists a collection of states $(y_1, \dots, y_n) \in \mathfrak{T}_n$ such that

$$\pi_{\mathfrak{m}_1} = y_{\sigma_1} \in \mathcal{S}_{\sigma_1}, \dots, \pi_{\mathfrak{m}_n} = y_{\sigma_n} \in \mathcal{S}_{\sigma_n}.$$

Now, every $\boldsymbol{\pi} \in \Pi_0$ can be written as a composition $\boldsymbol{\pi} = \boldsymbol{\pi}_\sigma \circ \tilde{\boldsymbol{\pi}}$ where $\tilde{\boldsymbol{\pi}} \in \Pi_0$ such that $\tilde{\pi}_{\mathfrak{m}_i} \in \mathcal{S}_i$ and where $\boldsymbol{\pi}_\sigma$ only acts on the states $y_1 = \tilde{\pi}_{\mathfrak{m}_1}, \dots, y_n = \tilde{\pi}_{\mathfrak{m}_n}$ according to the permutation $\boldsymbol{\sigma}$. So let $\sigma(y_1, \dots, y_n)$ denote a cycle representation of $\boldsymbol{\pi}_\sigma$. With the notations introduced in the previous lemmas we then can write $\boldsymbol{\pi}$ as

$$\boldsymbol{\pi} = \sigma(y_1, \dots, y_n)(\mathfrak{m}_1\, y_1 \cdots T^{k_1 - 1}(y_1)) \cdot \ldots \cdot (\mathfrak{m}_n\, y_n \cdots T^{k_n - 1}(y_n)).$$

Keep in mind that the numbers k_i also depend on y_i and that one always has $T^{k_i}(y_i) = \mathfrak{m}_i$. Since f_0 does not take $\mathfrak{m}_1, \dots, \mathfrak{m}_n$ into account we have $f_0(\boldsymbol{\pi}) = f_0(\tilde{\boldsymbol{\pi}})$. Combining that

with the equality $\operatorname{sgn}(\boldsymbol{\sigma}) = \operatorname{sgn}(\boldsymbol{\pi}_\sigma)$ and Lemma 4.2.9 yields

$$\operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} = \operatorname{sgn}(\boldsymbol{\pi}_\sigma)\operatorname{sgn}(\tilde{\boldsymbol{\pi}})(-1)^{f_0(\tilde{\boldsymbol{\pi}})} = \operatorname{sgn}(\boldsymbol{\sigma})(-1)^{N-n}.$$

□

This allows us to rewrite the first summand in (4.2.7):

$$\begin{aligned}
\sum_{\boldsymbol{\pi}\in\Pi_0} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} \prod_{i=1}^{n} a^{\varepsilon}_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) &= \sum_{\boldsymbol{\sigma}\in S_n} \sum_{\boldsymbol{\pi}\in\Pi(\boldsymbol{\sigma})} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} \prod_{i=1}^{n} a^{\varepsilon}_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) \\
&= \sum_{\boldsymbol{\sigma}\in S_n} \operatorname{sgn}(\boldsymbol{\sigma})(-1)^{N-n} \sum_{\boldsymbol{\pi}\in\Pi(\boldsymbol{\sigma})} \prod_{i=1}^{n} a^{\varepsilon}_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) \\
&= (-1)^{N-n} \sum_{\boldsymbol{\sigma}\in S_n} \operatorname{sgn}(\boldsymbol{\sigma}) \sum_{y_1\in\mathcal{S}_{\sigma_1}} \cdots \sum_{y_n\in\mathcal{S}_{\sigma_n}} \prod_{i=1}^{n} a^{\varepsilon}_{\mathfrak{m}_i,y_i}(\lambda) \\
&= (-1)^{N-n} \sum_{\boldsymbol{\sigma}\in S_n} \operatorname{sgn}(\boldsymbol{\sigma}) \prod_{i=1}^{n} \sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y}(\lambda). \qquad (4.2.8)
\end{aligned}$$

Here we can see why we also need the term $(-1)^{N-n}$ in the normalization factor for P^ε in Proposition 4.2.3.

The next lemma provides us with the crucial estimate.

Lemma 4.2.10. *There exist constants $\varepsilon_0 > 0$ and $\tilde{C} = \tilde{C}(h, \delta, R, \alpha, n)$ such that for every $\lambda \in \mathbb{C}$ and every $0 < \varepsilon < \varepsilon_0$*

$$\left| \sum_{\boldsymbol{\sigma}\in S_n} \operatorname{sgn}(\boldsymbol{\sigma}) \prod_{i=1}^{n} \sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y}(\lambda) - \sum_{\boldsymbol{\sigma}\in S_n} \operatorname{sgn}(\boldsymbol{\sigma}) \prod_{i=1}^{n} a^{\mathbf{Q}}_{i,\sigma_i}(\lambda) \right| \leq \tilde{C} \max\left\{1; (\hat{q} + |\lambda|)^{n-1}\right\} \varepsilon^{\alpha}. \qquad (4.2.9)$$

Proof. Abbreviate the left hand side of (4.2.9) with (LHS). We apply Lemma A.4 with

$c_i = \sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y}(\lambda) - a^{\mathbf{Q}}_{i,\sigma_i}(\lambda)$ and $d_i = a^{\mathbf{Q}}_{i,\sigma_i}(\lambda)$ and obtain

$$
\begin{aligned}
(LHS) &\leq \sum_{\boldsymbol{\sigma}\in S_n} \left| \prod_{i=1}^{n} \left(\sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y}(\lambda) - a^{\mathbf{Q}}_{i,\sigma_i}(\lambda) + a^{\mathbf{Q}}_{i,\sigma_i}(\lambda) \right) - \prod_{i=1}^{n} a^{\mathbf{Q}}_{i,\sigma_i}(\lambda) \right| \\
&= \sum_{\boldsymbol{\sigma}\in S_n} \left| \prod_{i=1}^{n} c_i + \sum_{i=1}^{n-1} \sum_{1\leq j_1<\ldots<j_i\leq n} \prod_{k\in\{j_1,\ldots,j_i\}} c_k \prod_{\substack{l\in\{1,\ldots,n\} \\ l\notin\{j_1,\ldots,j_i\}}} d_l \right| \\
&\leq \sum_{\boldsymbol{\sigma}\in S_n} \prod_{i=1}^{n} \left| \sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y}(\lambda) - a^{\mathbf{Q}}_{i,\sigma_i}(\lambda) \right| + \\
&\quad \sum_{\boldsymbol{\sigma}\in S_n} \sum_{i=1}^{n-1} \sum_{1\leq j_1<\ldots<j_i\leq n} \underbrace{\prod_{k\in\{j_1,\ldots,j_i\}} \left| \sum_{y\in\mathcal{S}_{\sigma_k}} a^{\varepsilon}_{\mathfrak{m}_k,y}(\lambda) - a^{\mathbf{Q}}_{k,\sigma_k}(\lambda) \right| \prod_{\substack{l\in\{1,\ldots,n\} \\ l\notin\{j_1,\ldots,j_i\}}} \left| a^{\mathbf{Q}}_{l,\sigma_l}(\lambda) \right|}_{(*)}.
\end{aligned}
\tag{4.2.10}
$$

Furthermore, since each row sum of $\mathbf{P}^{\varepsilon} - \mathbf{I}_N$ is 0 we have for every $x \in \mathcal{S}$

$$
\sum_{y\in\mathcal{S}_i} \left(\mathbf{P}^{\varepsilon} - \mathbf{I}_N\right)_{x,y} = -\sum_{\substack{j=1 \\ j\neq i}}^{n} \sum_{y\in\mathcal{S}_j} \left(\mathbf{P}^{\varepsilon} - \mathbf{I}_N\right)_{x,y}
$$

Now, for $\boldsymbol{\sigma} \in S_n$ define $F(\boldsymbol{\sigma}) \coloneqq \{1 \leq i \leq n |\ \sigma_i = i\}$ as the set of all fixed points of $\boldsymbol{\sigma}$. Then, with the help of Lemma 4.1.2 we derive

$$
\begin{aligned}
&\sum_{\boldsymbol{\sigma}\in S_n} \prod_{i=1}^{n} \left| \sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y}(\lambda) - a^{\mathbf{Q}}_{i,\sigma_i}(\lambda) \right| \\
&\quad = \sum_{\boldsymbol{\sigma}\in S_n} \left[\prod_{\substack{i=1 \\ i\notin F(\boldsymbol{\sigma})}}^{n} \left| \sum_{y\in\mathcal{S}_{\sigma_i}} a^{\varepsilon}_{\mathfrak{m}_i,y} - a^{\mathbf{Q}}_{i,\sigma_i} \right| \prod_{\substack{i=1 \\ i\in F(\boldsymbol{\sigma})}}^{n} \left| -\sum_{\substack{j=1 \\ j\neq i}}^{n} \sum_{y\in\mathcal{S}_j} a^{\varepsilon}_{\mathfrak{m}_i,y} - \lambda - \left(-\sum_{\substack{j=1 \\ j\neq i}}^{n} q_{i,j} - \lambda \right) \right| \right] \\
&\quad = \sum_{\boldsymbol{\sigma}\in S_n} \left[\prod_{\substack{i=1 \\ i\notin F(\boldsymbol{\sigma})}}^{n} \left| \sum_{y\in\mathcal{S}_{\sigma_i}} \frac{c(\alpha,h)}{\varepsilon^{\alpha}} p^{\varepsilon}_{\mathfrak{m}_i,y} - q_{i,\sigma_i} \right| \prod_{\substack{i=1 \\ i\in F(\boldsymbol{\sigma})}}^{n} \left| -\sum_{\substack{j=1 \\ j\neq i}}^{n} \sum_{y\in\mathcal{S}_j} \frac{c(\alpha,h)}{\varepsilon^{\alpha}} p^{\varepsilon}_{\mathfrak{m}_i,y} + \sum_{\substack{j=1 \\ j\neq i}}^{n} q_{i,j} \right| \right] \\
&\quad \leq \sum_{\boldsymbol{\sigma}\in S_n} \left(C\varepsilon^{\alpha}\right)^{n-|F(\boldsymbol{\sigma})|} \left(nC\varepsilon^{\alpha}\right)^{|F(\boldsymbol{\sigma})|} \\
&\quad \leq n! \left(nC\varepsilon^{\alpha}\right)^{n}.
\end{aligned}
$$

Similarly one can estimate the expression $(*)$ in (4.2.10) as follows:

$$\begin{aligned}
(*) \leq & \prod_{\substack{k\in\{j_1,\dots,j_i\}\\ k\notin F(\boldsymbol{\sigma})}} \left| \sum_{y\in\mathcal{S}_k} \frac{c(\alpha,h)}{\varepsilon^\alpha} p^\varepsilon_{\mathfrak{m}_k,y} - q_{k,\sigma_k} \right| \prod_{\substack{l\in\{1,\dots,n\}\\ l\notin\{j_1,\dots,j_i\}\\ l\notin F(\boldsymbol{\sigma})}} |q_{l,\sigma_l}| \\
& \cdot \prod_{\substack{k\in\{j_1,\dots,j_i\}\\ k\in F(\boldsymbol{\sigma})}} \left| -\sum_{s\neq k}\sum_{y\in\mathcal{S}_s} \frac{c(\alpha,h)}{\varepsilon^\alpha} p^\varepsilon_{\mathfrak{m}_k,y} + \sum_{s\neq k} q_{k,s} \right| \prod_{\substack{l\in\{1,\dots,n\}\\ l\notin\{j_1,\dots,j_i\}\\ l\notin F(\boldsymbol{\sigma})}} |q_{l,l}-\lambda| \\
\leq & \, (C\varepsilon^\alpha)^{|\{k\in\{j_1,\dots,j_i\}|\ \sigma_k\neq k\}|} \, (nC\varepsilon^\alpha)^{|\{k\in\{j_1,\dots,j_i\}|\ \sigma_k=k\}|} \\
& \cdot \hat{q}^{\{k\in\{1,\dots,n\}\setminus\{j_1,\dots,j_i\}|\ \sigma_k\neq k\}} \, (\hat{q}+|\lambda|)^{\{k\in\{1,\dots,n\}\setminus\{j_1,\dots,j_i\}|\ \sigma_k=k\}} \\
\leq & \, (nC\varepsilon^\alpha)^i (\hat{q}+|\lambda|)^{n-i}.
\end{aligned}$$

We finish the proof by continuing in (4.2.10):

$$(LHS) \leq n!\,(nC\varepsilon^\alpha)^n + n!\sum_{i=1}^{n-1}\binom{n}{i}(nC\varepsilon^\alpha)^i(\hat{q}+|\lambda|)^{n-i} \leq \tilde{C}\max_{1\leq k\leq n-1}(\hat{q}+|\lambda|)^k\varepsilon^\alpha.$$

□

We are now in the position to complete the proof of Proposition 4.2.3.

Proof of Proposition 4.2.3: First, we combine the equations (4.2.7) and (4.2.8) to obtain

$$\begin{aligned}
\left(-\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n} P^\varepsilon(\lambda) &= \sum_{\boldsymbol{\sigma}\in S_n} \operatorname{sgn}(\boldsymbol{\sigma}) \prod_{i=1}^{n} \sum_{y\in\mathcal{S}_{\sigma_i}} a^\varepsilon_{\mathfrak{m}_i,y} \\
&+ (-1)^{N-n} \sum_{\boldsymbol{\pi}\in\Pi_0} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_0(\boldsymbol{\pi})} \left(\sum_{l=0}^{N-n-1} \beta^\varepsilon_{l,0,\boldsymbol{\pi}}(\lambda) \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l} \right) \prod_{i=1}^{n} a^\varepsilon_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) \\
&+ (-1)^{N-n} \sum_{p=1}^{N-n} \sum_{\boldsymbol{\pi}\in\Pi_p} \operatorname{sgn}(\boldsymbol{\pi})(-1)^{f_p(\boldsymbol{\pi})} \left(\sum_{l=0}^{N-n-p} \beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda) \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l} \right) \prod_{x\in\mathcal{S}^3(\boldsymbol{\pi})} a^\varepsilon_{x,\pi_x}.
\end{aligned}$$

An application of Lemma 4.2.10 then yields

$$\begin{aligned}\left|\left(-\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n} P^\varepsilon(\lambda) - P^{\boldsymbol{Q}}(\lambda)\right| &\le \tilde{C} \max_{1\le k\le n-1} (\hat{q}+|\lambda|)^k \varepsilon^\alpha \\ &+ \sum_{\boldsymbol{\pi}\in\Pi_0} \left| \sum_{l=0}^{N-n-1} \beta^\varepsilon_{l,0,\boldsymbol{\pi}}(\lambda) \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l} \prod_{i=1}^{n} a^\varepsilon_{\mathfrak{m}_i,\pi_{\mathfrak{m}_i}}(\lambda) \right| \\ &+ \sum_{p=1}^{N-n} \sum_{\boldsymbol{\pi}\in\Pi_p} \left| \sum_{l=0}^{N-n-p} \beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda) \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l} \prod_{x\in\mathcal{S}^3(\boldsymbol{\pi})} a^\varepsilon_{x,\pi_x} \right|. \end{aligned} \tag{4.2.11}$$

We want to estimate the second and third summand. For this let us use Lemma 4.2.7 to renumber the states of $\mathcal{S}^1(\boldsymbol{\pi})$ and $\mathcal{S}^2(\boldsymbol{\pi})$ such that $\mathcal{S}^1(\boldsymbol{\pi}) = \left\{x_1,\dots,x_{f_p(\boldsymbol{\pi})}\right\}$ and $\mathcal{S}^2(\boldsymbol{\pi}) = \left\{x_{f_p(\boldsymbol{\pi})+1},\dots,x_{N-n-p}\right\}$. Then, from (4.2.6) it is easy to see that the numbers $\beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda)$ are polynomials in λ of degree $|\mathcal{S}^1(\boldsymbol{\pi})| = f_p(\boldsymbol{\pi})$. Expanding yields

$$\beta^\varepsilon_{0,p,\boldsymbol{\pi}}(\lambda) = \prod_{\nu=1}^{f_p(\boldsymbol{\pi})} (\lambda - \tilde{a}^\varepsilon_{x_\nu,x_\nu}) \prod_{\nu=f_p(\boldsymbol{\pi})+1}^{N-n-p} (-\tilde{a}^\varepsilon_{x_\nu,x_\nu})$$

as well as for $1 \le l \le N-n-p$

$$\beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda) = \sum_{1\le\mu_1<\dots<\mu_l\le N-n-p} \prod_{\substack{\nu=1\\ \nu\notin\{\mu_1,\dots,\mu_l\}}}^{f_p(\boldsymbol{\pi})} (\lambda - a^\varepsilon_{x_\nu,x_\nu}) \prod_{\substack{\nu=f_p(\boldsymbol{\pi})+1\\ \nu\notin\{\mu_1,\dots,\mu_l\}}}^{N-n-p} (-\tilde{a}^\varepsilon_{x_\nu,x_\nu}).$$

By Lemma 4.1.1 and Proposition 2.2.1 we then can find a constant $\overline{c} = \overline{c}(\alpha,h,\delta,R)$ such that for $0 < \varepsilon < \varepsilon_0$

$$\begin{aligned}\left|\beta^\varepsilon_{0,p,\boldsymbol{\pi}}(\lambda)\right| = \prod_{\nu=1}^{f_p(\boldsymbol{\pi})} \left|\lambda - \tilde{a}^\varepsilon_{x_\nu,x_\nu}\right| \prod_{\nu=f_p(\boldsymbol{\pi})+1}^{N-n-p} \left|\tilde{a}^\varepsilon_{x_\nu,x_\nu}\right| &= \prod_{\nu=1}^{f_p(\boldsymbol{\pi})} \left|\lambda + \sum_{\substack{\kappa=1\\ \kappa\ne\nu}}^{N} \tilde{a}^\varepsilon_{x_\kappa,x_\kappa}\right| \prod_{\nu=f_p(\boldsymbol{\pi})+1}^{N-n-p} \left|\tilde{a}^\varepsilon_{x_\nu,x_\nu}\right| \\ &\le \left(\prod_{\nu=1}^{f_p(\boldsymbol{\pi})} (|\lambda| + N\overline{c})\right) (\overline{c})^{N-n-p-f_p(\boldsymbol{\pi})} \\ &\le (|\lambda| + N\overline{c})^{f_p(\boldsymbol{\pi})} (\overline{c})^{N-n-p-f_p(\boldsymbol{\pi})} \\ &\le (|\lambda| + N\overline{c})^{N-n-p}\end{aligned}$$

and similarly

$$\left|\beta^\varepsilon_{l,p,\boldsymbol{\pi}}(\lambda)\right| \le N!(|\lambda| + N\overline{c})^{N-n-p-l}, \quad 1 \le l \le N-n-p.$$

Also, note that for $\boldsymbol{\pi} \in \Pi_p$ we have $|\mathcal{S}^3(\boldsymbol{\pi})| = n + p$. Then

$$\begin{aligned}\left|\prod_{x\in\mathcal{S}^3(\boldsymbol{\pi})} a^\varepsilon_{x,\pi_x}(\lambda)\right| &= \prod_{\substack{x\in\mathcal{S}^3(\boldsymbol{\pi})\\ \pi_x=x}} \left|\lambda + \tilde{a}^\varepsilon_{x,x}\right| \prod_{\substack{x\in\mathcal{S}^3(\boldsymbol{\pi})\\ \pi_x\neq x}} \left|a^\varepsilon_{x,\pi_x}\right| \\ &\leq (|\lambda| + N\bar{c})^{n+p}\,.\end{aligned}$$

Finally, continuing in (4.2.11) yields for ε sufficiently small

$$\begin{aligned}&\left|\left(-\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n} P^\varepsilon(\lambda) - P^{\boldsymbol{Q}}(\lambda)\right| \\ &\qquad\leq \tilde{C} \max_{1\leq k\leq n-1} (\hat{q} + |\lambda|)^k \varepsilon^\alpha \\ &\qquad\quad + (|\lambda| + N\bar{c})^n \sum_{\boldsymbol{\pi}\in\Pi_0} \sum_{l=0}^{N-n-1} N!(|\lambda| + N\bar{c})^{N-n-l} \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l} \\ &\qquad\quad + \sum_{p=1}^{N-n} (|\lambda| + N\bar{c})^{n+p} \sum_{\boldsymbol{\pi}\in\Pi_p} \sum_{l=0}^{N-n-p} N!(|\lambda| + N\bar{c})^{N-n-p-l} \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^{N-n-l} \\ &\qquad\leq \tilde{C} \max_{1\leq k\leq n-1} (\hat{q} + |\lambda|)^k \varepsilon^\alpha \\ &\qquad\quad + (|\lambda| + N\bar{c})^n (N!)^2 \left[\sum_{l=0}^{\infty} (|\lambda| + N\bar{c})^l \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^l - 1\right] \\ &\qquad\quad + (|\lambda| + N\bar{c})^n (N-n)(N!)^2 \left[\sum_{l=0}^{\infty} (|\lambda| + N\bar{c})^l \left(\frac{\varepsilon^\alpha}{c(\alpha,h)}\right)^l - 1\right] \\ &\qquad\leq \tilde{C} \max_{1\leq k\leq n-1} (\hat{q} + |\lambda|)^k \varepsilon^\alpha \\ &\qquad\quad + (|\lambda| + N\bar{c})^{n+1} (N!)^2 \frac{2}{c(\alpha,h)} \varepsilon^\alpha \\ &\qquad\quad + (|\lambda| + N\bar{c})^{n+1} (N-n)(N!)^2 \frac{2}{c(\alpha,h)} \varepsilon^\alpha.\end{aligned}$$

That concludes the proof.

□

Note that we did not use any a priori information about the eigenvalues of $\mathbf{Q}$ other than the fact that one of them is 0. However, from now on we will make the following additional assumption:

(E) All eigenvalues of $\mathbf{Q}$ are real and simple.

We strongly believe that this assumption always holds. This is supported by numerous computer simulations of such a matrix $\mathbf{Q}$, i.e. with entries of the form (3.0.1) which we randomly generated. The construction of such a matrix is relatively easy. For example, generate $2n-1$ random real numbers $x_1, \dots, x_{2n-1}$ and define

$$u(x) \coloneqq (x - x_1) \cdots (x - x_{2n-1})$$

and U as the antiderivative of u. Then these real numbers are the local extrema of U and hence the elements of the set $\mathcal{M}$. The spectrum always consisted of real numbers so we think it should be possible to prove that it is always true. However, since $\mathbf{Q}$ is not necessarily symmetric this problem seems to be quite complicated.

4.3 Eigenvectors

Let $\lambda_1^\varepsilon, \dots, \lambda_n^\varepsilon$ be the eigenvalues from Theorem 4.2.2 (i). Denote by $\tilde{\lambda}_1, \tilde{\lambda}_2^\varepsilon, \dots, \tilde{\lambda}_n^\varepsilon$ the eigenvalues of the (unscaled) matrix $\mathbf{I}_N - \mathbf{P}^\varepsilon$. From Theorem 4.2.2 it follows that

$$\tilde{\lambda}_1^\varepsilon = 0 \quad \text{and} \quad \tilde{\lambda}_i^\varepsilon = O(\varepsilon^\alpha), \quad 2 \le i \le n. \tag{4.3.1}$$

Obviously, in regard to the investigation of eigenvectors it does not matter if we consider the scaled matrix $\mathbf{Q}^\varepsilon$, and with that the scaled eigenvalues from Theorem 4.2.2, or the unscaled one. However, we will follow closely the ideas presented in the series of papers [BEGK2], [Eck] and therefore it is more convenient to work with $\tilde{\lambda}_i^\varepsilon$ instead of λ_i^ε. Note that we also consider $\mathbf{I}_N - \mathbf{P}^\varepsilon$ instead of $\mathbf{P}^\varepsilon - \mathbf{I}_N$ which in view of Proposition 4.2.1 implies

$$\operatorname{Re}\left(\tilde{\lambda}_i^\varepsilon\right) \ge 0, \quad 1 \le i \le n.$$

Let us state the main result of this section. Denote by $\boldsymbol{\delta}^j = (\delta_x^j)_{x \in \mathcal{S}}$, $1 \le j \le n$, the indicator vector of $\mathcal{S}_j$, i.e. $\delta_x^j = 1$ for $x \in \mathcal{S}_j$ and 0 elsewhere.
Also recall that the eigenvector $\boldsymbol{\psi}^{\varepsilon,1} = (\psi_x^{\varepsilon,1})_{x \in \mathcal{S}}$ corresponding to $\lambda_1^\varepsilon = 0$ is constant.

Theorem 4.3.1. *Fix $2 \leq i \leq n$ and let $\boldsymbol{\psi}^{\varepsilon,i} = (\psi_x^{\varepsilon,i})_{x\in\mathcal{S}}$ be the right eigenvector associated with $\tilde{\lambda}_i^\varepsilon$ and normalized such that $\max_{1\leq j\leq n}\left|\psi_{\mathfrak{m}_j}^{\varepsilon,i}\right| = 1$. Moreover, let $\boldsymbol{\psi}^{\mathbf{Q},i} = (\psi_j^{\mathbf{Q},i})_{j=1}^n$ the right eigenvector of $\mathbf{Q}$ associated with $\lambda_i^{\mathbf{Q}}$ normalized such that $\max_{1\leq j\leq n}\left|\psi_j^{\mathbf{Q},i}\right| = 1$. Then, for every $x \in \mathcal{S}$, we have*

$$\left|\psi_x^{\varepsilon,i} - \sum_{j=1}^n \psi_{\mathfrak{m}_j}^{\mathbf{Q},i}\delta_x^j\right| \to 0, \quad \varepsilon \to 0. \tag{4.3.2}$$

The proof of this theorem needs some preparations. For a given subset $I \subset \mathcal{S}$ consider again the first entry time of I,

$$\sigma_I^\varepsilon \coloneqq \inf\left\{k \geq 0 |\ Z_k^\varepsilon(\cdot) \in I\right\},$$

and, for convenience, the first return time to I,

$$\tau_I^\varepsilon \coloneqq \inf\left\{k \geq 1 |\ Z_k^\varepsilon(\cdot) \in I\right\}.$$

We will now analyze certain Laplace transforms of these quantities. Define for $u \in \mathbb{C}$, $x \in \mathcal{S}$ and $I, J \subset \mathcal{S}$

$$G_{I,J}^{x,\varepsilon}(u) \coloneqq \mathbb{E}_x\left[\mathrm{e}^{u\tau_I^\varepsilon}\,\mathbf{1}_{\left\{\tau_I^\varepsilon \leq \tau_J^\varepsilon\right\}}\right] \tag{4.3.3}$$

and

$$K_{I,J}^{x,\varepsilon}(u) \coloneqq \mathbb{E}_x\left[\mathrm{e}^{u\sigma_I^\varepsilon}\,\mathbf{1}_{\left\{\sigma_I^\varepsilon \leq \sigma_J^\varepsilon\right\}}\right]. \tag{4.3.4}$$

Note that these expressions are finite for every $u \in \mathbb{C}$ such that $\mathrm{Re}(u) \leq 0$ but they can be infinite if the real part is positive. We will discuss the radius of convergence for certain cases later.

Since for an initial state $x \in \mathcal{S} \setminus I$ we have $\sigma_I^\varepsilon = \tau_I^\varepsilon$ it follows that

$$K_{I,J}^{x,\varepsilon}(u) = \begin{cases} G_{I,J}^{x,\varepsilon}(u), & x \notin I \cup J, \\ 1, & x \in I, \\ 0, & x \in J \setminus I. \end{cases} \tag{4.3.5}$$

In the following lemma we will prove another useful connection between these two functions.

Lemma 4.3.2. *Fix an arbitrary $x \in \mathcal{S}$, subsets $I, J \subset \mathcal{S}$ and $u \in \mathbb{C}$ such that $G^{x,\varepsilon}_{I,J}(u)$ is finite. Then*

$$\mathrm{e}^u \sum_{y\in\mathcal{S}} p^{\varepsilon}_{x,y} K^{y,\varepsilon}_{I,J}(u) = G^{x,\varepsilon}_{I,J}(u).$$

Proof. An application of the Markov property yields

$$\begin{aligned} G^{x,\varepsilon}_{I,J}(u) = \mathbb{E}_x\left[\mathrm{e}^{u\tau^{\varepsilon}_I}\mathbf{1}_{\{\tau^{\varepsilon}_I \le \tau^{\varepsilon}_J\}}\right] &= \sum_{y\in\mathcal{S}} \mathbb{E}_x\left[\mathrm{e}^{u(1+\sigma^{\varepsilon}_I)}\mathbf{1}_{\{1+\sigma^{\varepsilon}_I \le 1+\sigma^{\varepsilon}_J\}} | Z^{\varepsilon}_1 = y\right] \mathbb{P}_x(Z^{\varepsilon}_1 = y) \\ &= \mathrm{e}^u \sum_{y\in\mathcal{S}} p^{\varepsilon}_{x,y} K^{y,\varepsilon}_{I,J}(u). \end{aligned}$$

□

Now let us focus on the case $I, J \subset \mathcal{M}$. We are interested in the radius of convergence of these Laplace transforms in dependence on ε. For this let

$$u^{\varepsilon}_{\mathcal{M}} \coloneqq \sup\left\{u \ge 0 |\ G^{x,\varepsilon}_{\mathcal{M},\mathcal{M}}(u) = \mathbb{E}_x\,\mathrm{e}^{u\tau^{\varepsilon}_{\mathcal{M}}} < \infty \quad \text{for all } x \in \mathcal{S}\right\}.$$

Obviously, for $u \ge 0$ and $I \subset \mathcal{M}$ one always has $G^{x,\varepsilon}_{I,\mathcal{M}}(u) \le G^{x,\varepsilon}_{\mathcal{M},\mathcal{M}}(u)$ so we can immediately conclude $G^{x,\varepsilon}_{I,\mathcal{M}}(u) < \infty$ for every $u \in \mathbb{C}$ with $0 \le \mathrm{Re}(u) \le u^{\varepsilon}_{\mathcal{M}}$. The following proposition shows that, for arbitrary $u \in \mathbb{C}$, the convergence is always ensured as long as ε is chosen sufficiently small.

Proposition 4.3.3. *We have*

$$\lim_{\varepsilon\to 0} u^{\varepsilon}_{\mathcal{M}} = \infty.$$

Proof. Let again $(\xi_k)_{k\ge 1}$ be the sequences of independent and identically distributed random variables such that $\xi_1 \overset{d}{=} L_1$. Recall the definition of $T^{\max}$ as the maximum time the deterministic motion needs to get to $\mathcal{M}$ when starting somewhere in $\mathcal{S}$,

$$T^{\max} = \max_{x\in\mathcal{S}} \min\left\{k \ge 1 |\ Z^0_k(x) \in \mathcal{M}\right\}.$$

Now, consider $k \ge 1$ such that $T^{\max} + 1 \le k \le 2T^{\max}$. Then the occurrence of the event $\{\tau^{\varepsilon}_{\mathcal{M}} = k\}$ implies that a big jump must have happened at least once, i.e.

$$\{\tau^{\varepsilon}_{\mathcal{M}} = k\} \subset \left\{|\varepsilon h^{1/\alpha}\xi_l| > D \text{ for at least one } 1 \le l \le k\right\}.$$

Similarly, for $2T^{\max}+1 \le k \le 3T^{\max}$ the occurrence of $\{\tau^\varepsilon_{\mathcal{M}} = k\}$ implies that a big jump happened at least twice and so forth. Hence, for $\mu \in \mathbb{N}$ and $\mu T^{\max}+1 \le k \le (\mu+1)T^{\max}$ one has

$$\{\tau^\varepsilon_{\mathcal{M}} = k\} \subset \left\{ |\varepsilon h^{1/\alpha}\xi_l| > D \text{ for at least } \mu \text{ times } \{l_1,\dots,l_\mu\} \subset \{1,\dots,k\} \right\}. \tag{4.3.6}$$

Recalling that $p^\varepsilon := \mathbb{P}\left(|\varepsilon 1/\alpha \xi_1| > D\right)$. The event on the right hand side of (4.3.6) does not depend on x and therefore we find for all $x \in \mathcal{S}$ and $\mu \in \mathbb{N}$ and $\mu T^{\max}+1 \le k \le (\mu+1)T^{\max}$

$$\begin{aligned}
\mathbb{P}_x(\tau^\varepsilon_{\mathcal{M}} = k) &\le \mathbb{P}\left(|\varepsilon h^{1/\alpha}\xi_l| > D \text{ for at least } \mu \text{ times } \{l_1,\dots,l_\mu\} \subset \{1,\dots,k\} \right) \\
&\le \mathbb{P}\left(|\varepsilon h^{1/\alpha}\xi_l| > D \text{ for at least } \mu \text{ times } \{l_1,\dots,l_\mu\} \subset \{1,\dots,(\mu+1)T^{\max}\} \right) \\
&= \sum_{l=\mu}^{(\mu+1)T^{\max}} \binom{(\mu+1)T^{\max}}{l} (p^\varepsilon)^l (1-p^\varepsilon)^{(\mu+1)T^{\max}-l}.
\end{aligned}$$

Let us estimate the binomial coefficient. By applying Stirling's formula we find a constant $C > 0$ such that

$$\binom{(\mu+1)T^{\max}}{l} \le C\sqrt{\frac{(\mu+1)T^{\max}}{l((\mu+1)T^{\max}-l)}} \left(\frac{(\mu+1)T^{\max}}{(\mu+1)T^{\max}-l}\right)^{(\mu+1)T^{\max}-l}.$$

From $\mu \le l \le (\mu+1)T^{\max}$ it follows that l increases linearly with μ. Therefore, there exist constants $\tilde{C} > 0$ and $K > 0$, independent of μ and ε, such that

$$\binom{(\mu+1)T^{\max}}{l} \le \tilde{C}\sqrt{\frac{1}{\mu}}\, K^{(\mu+1)T^{\max}}.$$

We then can continue the estimate of $\mathbb{P}_x(\tau^\varepsilon_{\mathcal{M}} = k)$ as follows:

$$\mathbb{P}_x(\tau^\varepsilon_{\mathcal{M}} = k) \le \tilde{C}\left((\mu+1)T^{\max}-\mu+1\right)\sqrt{\frac{1}{\mu}}\, K^{(\mu+1)T^{\max}} (p^\varepsilon)^\mu$$

Then, setting $\hat{C} = \sum_{k=1}^{T^{\max}} \mathrm{e}^{u^\varepsilon k}\, \mathbb{P}_x(\tau^\varepsilon_{\mathcal{M}} = k)$ we have

$$\begin{aligned}
\mathbb{E}_x\, \mathrm{e}^{u^\varepsilon \tau^\varepsilon_{\mathcal{M}}} &= \sum_{k=1}^{\infty} \mathrm{e}^{u^\varepsilon k}\, \mathbb{P}_x(\tau^\varepsilon_{\mathcal{M}} = k) \\
&= \hat{C} + \sum_{\mu=1}^{\infty} \sum_{k=\mu T^{\max}+1}^{(\mu+1)T^{\max}} \mathrm{e}^{u^\varepsilon k}\, \mathbb{P}_x(\tau^\varepsilon_{\mathcal{M}} = k) \\
&\le \hat{C} + \tilde{C}\sum_{\mu=1}^{\infty} \mathrm{e}^{u^\varepsilon (\mu+1)T^{\max}}\, T^{\max}\left((\mu+1)T^{\max}-\mu+1\right)\sqrt{\frac{1}{\mu}}\, K^{(\mu+1)T^{\max}} (p^\varepsilon)^\mu
\end{aligned}$$

This series converges if $\mathrm{e}^{uT^{\max}} K^{T^{\max}} p^{\varepsilon} < 1$ or, equivalently, if $u > 0$ is chosen such that $u \leq \frac{1}{T^{\max}} \ln\left(\frac{1}{p^{\varepsilon}}\right) - \ln(K)$. The claim of the proposition follows from the fact that p^{ε} is of order $O(\varepsilon^{\alpha})$.

□

From this proposition it follows that for a fixed $u_0 > 0$ one can always choose ε small enough such that $u_0 \leq u^{\varepsilon}_{\mathcal{M}}$ and therefore we can immediately deduce the following corollary that will provide us with an approximation of $G^{x,\varepsilon}_{I,J}(u)$ with respect to u in a neighborhood of 0.

Corollary 4.3.4. *Fix $u_0 > 0$ and $\varepsilon_0 > 0$ such that $u_0 \leq u^{\varepsilon}_{\mathcal{M}}$ and $0 < \varepsilon < \varepsilon_0$. Moreover, let $x \in \mathcal{S}$ and $I \subseteq \mathcal{M}$. Then there exists a constant $L > 0$, which can be chosen independent of ε and x, such that for every $u \leq u_0$*

$$\left|G^{x,\varepsilon}_{I,\mathcal{M}}(u) - G^{x,\varepsilon}_{I,\mathcal{M}}(0)\right| \leq L \cdot |u|.$$

Proof. The choice of ε guarantees that $G^{x,\varepsilon}_{I,\mathcal{M}}(u)$ is finite for $u \leq u_0$. We then can use Proposition 1.2.3 and obtain

$$\left|G^{x,\varepsilon}_{I,\mathcal{M}}(u) - G^{x,\varepsilon}_{I,\mathcal{M}}(0)\right| \leq \sum_{n=1}^{\infty} \frac{1}{n!} \left.\frac{d^n G^{x,\varepsilon}_{I,\mathcal{M}}(u)}{du^n}\right|_{u=0} |u|^n.$$

We have to estimate the expressions $\left.\frac{d^n G^{x,\varepsilon}_{I,\mathcal{M}}(u)}{du^n}\right|_{u=0}$ in the limit of small ε. An easy calculation yields

$$\begin{aligned}
\left.\frac{d^n G^{x,\varepsilon}_{I,\mathcal{M}}(u)}{du^n}\right|_{u=0} &= \left.\frac{d^n \sum_{k=1}^{\infty} \mathrm{e}^{uk}\, \mathbb{P}_x(\tau^{\varepsilon}_I = k \leq \tau^{\varepsilon}_{\mathcal{M}})}{du^n}\right|_{u=0} \\
&= \left.\sum_{k=1}^{\infty} k^n\, \mathrm{e}^{uk}\, \mathbb{P}_x(\tau^{\varepsilon}_I = k \leq \tau^{\varepsilon}_{\mathcal{M}})\right|_{u=0} \\
&= \sum_{k=1}^{\infty} k^n \mathbb{P}_x(\tau^{\varepsilon}_I = k \leq \tau^{\varepsilon}_{\mathcal{M}}) \\
&\leq \sum_{k=1}^{\infty} k^n \mathbb{P}_x(k \leq \tau^{\varepsilon}_{\mathcal{M}}) \\
&= \sum_{k=1}^{\infty} k^n \sum_{l=k}^{\infty} \mathbb{P}_x(\tau^{\varepsilon}_{\mathcal{M}} = l).
\end{aligned}$$

Now we proceed with similar arguments and the same constants $\tilde{C}$ and K as in the proof of Proposition 4.3.3. For a given $k \geq 1$ let $\mu_k \in \mathbb{N}$ be defined by the relation $\mu_k T^{\max} + 1 \leq k \leq (\mu_k + 1)T^{\max}$. Then,

$$\begin{aligned}\sum_{k=1}^{\infty} k^n \sum_{l=k}^{\infty} \mathbb{P}_x(\tau_{\mathcal{M}}^\varepsilon = l) &\leq \sum_{k=1}^{\infty} k^n \sum_{\mu=\mu_k}^{\infty} \sum_{l=\mu T^{\max}+1}^{(\mu+1)T^{\max}} \mathbb{P}_x(\tau_{\mathcal{M}}^\varepsilon = l) \\ &\leq \sum_{k=1}^{\infty} k^n \sum_{\mu=\mu_k}^{\infty} \tilde{C}T^{\max}\left((\mu+1)T^{\max} - \mu + 1\right)\sqrt{\frac{1}{\mu}}\, K^{(\mu+1)T^{\max}}(p^\varepsilon)^\mu \end{aligned} \tag{4.3.7}$$

The term $\left((\mu+1)T^{\max} - \mu + 1\right)\sqrt{\frac{1}{\mu}}$ increases like $\sqrt{\mu}$. Therefore, there exists a constant $\overline{C} > 0$ such that we can estimate

$$\tilde{C}T^{\max}\left((\mu+1)T^{\max} - \mu + 1\right)\sqrt{\frac{1}{\mu}}\, K^{(\mu+1)T^{\max}}(p^\varepsilon)^\mu \leq \overline{C}\left(2K^{T^{\max}}p^\varepsilon\right)^\mu.$$

Continuing in (4.3.7) yields for sufficiently small ε

$$\begin{aligned}\sum_{k=1}^{\infty} k^n \sum_{l=k}^{\infty} \mathbb{P}_x(\tau_{\mathcal{M}}^\varepsilon = l) &\leq \overline{C} \sum_{k=1}^{\infty} k^n \sum_{\mu=\mu_k}^{\infty} \left(2K^{T^{\max}}p^\varepsilon\right)^\mu \\ &= \overline{C} \sum_{k=1}^{\infty} k^n \left(2K^{T^{\max}}p^\varepsilon\right)^{\mu_k} \sum_{\mu=0}^{\infty} \left(2K^{T^{\max}}p^\varepsilon\right)^\mu \\ &= \overline{C} \sum_{k=1}^{\infty} k^n \left(2K^{T^{\max}}p^\varepsilon\right)^{\mu_k} \frac{1}{1 - 2K^{T^{\max}}p^\varepsilon}.\end{aligned}$$

Let us abbreviate $r^\varepsilon := \left(2K^{T^{\max}}p^\varepsilon\right)^{\frac{1}{T^{\max}}}$. Since $\mu_k \leq \frac{k-1}{T^{\max}}$ one can find yet another constant $C^* > 0$ such that

$$\frac{\overline{C}}{1 - 2K^{T^{\max}}p^\varepsilon} \sum_{k=1}^{\infty} k^n \left(2K^{T^{\max}}p^\varepsilon\right)^{\mu_k} \leq C^* \sum_{k=0}^{\infty} k^n (r^\varepsilon)^k.$$

and therefore

$$\begin{aligned}\left|G_{I,\mathcal{M}}^{x,\varepsilon}(u) - G_{I,\mathcal{M}}^{x,\varepsilon}(0)\right| &\leq C^* \sum_{n=1}^{\infty} \frac{|u|^n}{n!} \sum_{k=0}^{\infty} k^n (r^\varepsilon)^k = C^* \sum_{k=0}^{\infty} (r^\varepsilon)^k (\mathrm{e}^{k|u|} - 1) \\ &= C^* \frac{r^\varepsilon(\mathrm{e}^{|u|} - 1)}{(1 - r^\varepsilon)(1 - r^\varepsilon\,\mathrm{e}^{|u|})} \leq L|u|\end{aligned}$$

with a certain constant $L > 0$. □

The following proposition is a special case of Lemma 4.1 in [Eck] where the interested reader can also find a proof.

Proposition 4.3.5. *Let $\tilde{\lambda}^\varepsilon$ be an eigenvalue of $\mathbf{I}_N - \mathbf{P}^\varepsilon$, define u^ε by $\tilde{\lambda}^\varepsilon = 1 - \mathrm{e}^{-u^\varepsilon}$ and assume that that ε is small enough such that $0 < \mathrm{Re}(u^\varepsilon) < u^\varepsilon_{\mathcal{M}}$. Moreover, normalize the corresponding eigenvector $\boldsymbol{\psi}^{\boldsymbol{\varepsilon}} = (\psi^\varepsilon_x)_{x\in\mathcal{S}}$ such that $\psi^\varepsilon_{\mathfrak{m}_1} = 1$. Then*

$$\psi^\varepsilon_x = \sum_{j=1}^n \psi^\varepsilon_{\mathfrak{m}_j} K^{x,\varepsilon}_{\mathfrak{m}_j,\mathcal{M}}(u^\varepsilon), \quad x \in \mathcal{S}. \tag{4.3.8}$$

Now we can finish this section with the proof of the theorem.

Proof of Theorem 4.3.1: Let u^ε_i be defined by $\tilde{\lambda}^\varepsilon_i = 1 - \mathrm{e}^{u^\varepsilon_i}$. By (4.3.1) it follows that

$$\lim_{\varepsilon\to 0} u^\varepsilon_i = 0.$$

Now, fix $x \in \mathcal{S}$ and let $\hat{j}(x) \in \{1,\dots,n\}$ be the number uniquely determined by the relation $x \in \Omega_{\hat{j}(x)}$. Applying the relation (4.3.5) as well as Proposition 4.3.5 and Corollary 4.3.4 we derive

$$\begin{aligned}
\psi^{\varepsilon,i}_x &= \psi^{\varepsilon,i}_{\mathfrak{m}_{\hat{j}(x)}} K^{x,\varepsilon}_{\mathfrak{m}_{\hat{j}(x)},\mathcal{M}}(u^\varepsilon_i) + \sum_{j\neq\hat{j}(x)} \psi^{\varepsilon,i}_{\mathfrak{m}_j} K^{x,\varepsilon}_{\mathfrak{m}_j,\mathcal{M}}(u^\varepsilon_i) \\
&= \psi^{\varepsilon,i}_{\mathfrak{m}_{\hat{j}(x)}} K^{x,\varepsilon}_{\mathfrak{m}_{\hat{j}(x)},\mathcal{M}}(0) + \sum_{j\neq\hat{j}(x)} \psi^{\varepsilon,i}_{\mathfrak{m}_j} K^{x,\varepsilon}_{\mathfrak{m}_j,\mathcal{M}}(0) + O(u^\varepsilon_i) \\
&= \psi^{\varepsilon,i}_{\mathfrak{m}_{\hat{j}(x)}} \mathbb{P}_x\left(\sigma^\varepsilon_{\mathfrak{m}_{\hat{j}(x)}} \le \sigma^\varepsilon_{\mathcal{M}}\right) + \sum_{j\neq\hat{j}(x)} \psi^{\varepsilon,i}_{\mathfrak{m}_j} \mathbb{P}_x\left(\sigma^\varepsilon_{\mathfrak{m}_j} \le \sigma^\varepsilon_{\mathcal{M}}\right) + O(u^\varepsilon_i)
\end{aligned} \tag{4.3.9}$$

while the constant that appears in the Landau symbol can be chosen independently of ε (see Corollary 4.3.4). But since $x \in \Omega_{\hat{j}(x)}$ we also have

$$\lim_{\varepsilon\to 0} \mathbb{P}_x\left(\sigma^\varepsilon_{\mathfrak{m}_{\hat{j}(x)}} \le \sigma^\varepsilon_{\mathcal{M}}\right) = 1 \quad \text{as well as} \quad \lim_{\varepsilon\to 0} \mathbb{P}_x\left(\sigma^\varepsilon_{\mathfrak{m}_j} \le \sigma^\varepsilon_{\mathcal{M}}\right) = 0, \quad j \neq \hat{j}(x). \tag{4.3.10}$$

That is $\boldsymbol{\psi}^{\varepsilon,i} = \sum_{j=1}^n \psi^{\varepsilon,i}_{\mathfrak{m}_j} \mathbf{w}^{\varepsilon,j} + \mathbf{b}^\varepsilon$ where $w^{\varepsilon,j}_x := \mathbb{P}_x\left(\sigma^\varepsilon_{\mathfrak{m}_j} \le \sigma^\varepsilon_{\mathcal{M}}\right)$ and $\max_{x\in\mathcal{S}} |b^\varepsilon_x| = O(\varepsilon^\alpha)$. Using that representation in the eigenvalue equation $\mathbf{A}^\varepsilon(\lambda^\varepsilon_i)\boldsymbol{\psi}^{\varepsilon,i} = 0$ (for the definition of $\mathbf{A}^\varepsilon(\lambda)$ see (4.0.2) yields

$$\sum_{j=1}^n \psi^{\varepsilon,i}_{\mathfrak{m}_j} \mathbf{A}^\varepsilon(\lambda^\varepsilon_i)\mathbf{w}^{\varepsilon,j} = -\mathbf{A}^\varepsilon(\lambda^\varepsilon_i)\mathbf{b}^\varepsilon.$$

From this linear system consider the n equations indexed by $\mathfrak{m}_1, \dots, \mathfrak{m}_n$:

$$\sum_{j=1}^{n} \psi_{\mathfrak{m}_j}^{\varepsilon,i} \sum_{y \in \mathcal{S}} \left(\mathbf{A}^\varepsilon(\lambda_i^\varepsilon)\right)_{\mathfrak{m}_k,y} w_y^{\varepsilon,j} = -\sum_{y \in \mathcal{S}} \left(\mathbf{A}^\varepsilon(\lambda_i^\varepsilon)\right)_{\mathfrak{m}_k,y} b_y^\varepsilon, \quad 1 \le k \le n. \tag{4.3.11}$$

Abbreviate $\mathbf{A}^{\mathbf{Q}}(\lambda) \coloneqq \mathbf{Q} - \lambda \mathbf{I}_n$ and recall the following results:

(a) $\lim_{\varepsilon \to 0} \sum_{y \in \mathcal{S}_l} \left(\mathbf{A}^\varepsilon(\lambda)\right)_{\mathfrak{m}_k,y} = \left(\mathbf{A}^{\mathbf{Q}}(\lambda)\right)_{k,l}$ (see Lemma 4.1.2),

(b) $\lim_{\varepsilon \to 0} \lambda_i^\varepsilon = \lambda_i^{\mathbf{Q}}$ for $1 \le i \le n$ (see Theorem 4.2.2 (i)).

Using this and (4.3.10) as well as the fact that $\psi_{\mathfrak{m}_j}^{\varepsilon,i}$ is bounded with respect to ε due to the normalization one can see that there is a vector $\boldsymbol{\rho}^\varepsilon = (\rho_i^\varepsilon)_{i=1}^n$ such that $|\rho_i^\varepsilon| \to 0$, $1 \le i \le n$, and (4.3.11) can be written as

$$\mathbf{A}^{\mathbf{Q}}(\lambda_i^{\mathbf{Q}}) \psi_{\mathfrak{m}_k}^{\varepsilon,i} = \rho_k^\varepsilon, \quad 1 \le k \le n. \tag{4.3.12}$$

Finally, let us recall that the eigenspace of $\lambda_i^{\mathbf{Q}}$ is one dimensional (see assumption (**E**) at the end of Section 4.2). This allows us to prove the statement of Lemma 2.3.3 if we replace $\mathbf{Q}^T$ with $\mathbf{Q}(\lambda_i^{\mathbf{Q}})$ and $\boldsymbol{\pi}^{\mathbf{Q}}$ with $\boldsymbol{\psi}^{\mathbf{Q},i}$. In fact, Lemma 2.3.3 essentially corresponds to the special case $\lambda_1^{\mathbf{Q}} = 0$. Therefore, we can conclude from (4.3.12) that

$$\lim_{\varepsilon \to 0} \psi_{\mathfrak{m}_k}^{\varepsilon,i} = \psi_k^{\mathbf{Q},i}, \quad 1 \le k \le n.$$

Combining that with 4.3.9 finishes the proof.

□

5 The Relation of the Eigenvalues of the Generator and a Discrete Poisson Problem

In this chapter we want to present a method that gives access to the eigenvalues in Theorem 4.2.2 (i) while using the information on mean first transition times from Chapter 3.
As mentioned in the introduction the following inhomogeneous linear system of equations plays a crucial role in establishing this connection. Define for $1 \leq j \leq n$ the vector $\boldsymbol{\delta}^j = (\delta_x^j)_{x\in\mathcal{S}}$ by

$$\delta_x^j = \begin{cases} 1, & x \in \mathcal{S}_j, \\ 0, & x \notin \mathcal{S}_j. \end{cases}$$

Moreover, let $\boldsymbol{\pi}^\varepsilon$ denote the stationary distribution of Z^ε and let us abbreviate $\boldsymbol{\pi}^\varepsilon(\mathcal{S}_j) \coloneqq \sum_{x\in\mathcal{S}_j} \pi_x^\varepsilon$. Consider the system of equations

$$(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\boldsymbol{\tau}^{\varepsilon,j} = \boldsymbol{\pi}^\varepsilon(\mathcal{S}_j)\mathbf{1} - \boldsymbol{\delta}^j. \tag{5.0.1}$$

Since this system can be viewed as a discrete version of a Poisson type linear system of differential equations we will, from now on, refer to it as a *discrete Poisson problem.* In the Gaussian case Kolokoltsov and Makarov constructed in [KoMa] (see also [Ko2], Chapter 8, Section 2) a matrix of which the elements are the averages over the potential wells of the solutions to (5.0.1) with respect to the stationary measure $\boldsymbol{\pi}^\varepsilon$. So define a matrix $\mathbf{G}^\varepsilon = (g_{i,j}^\varepsilon)_{i,j=1}^n$ by

$$g_{i,j}^\varepsilon \coloneqq \frac{1}{\pi^\varepsilon(\mathcal{S}_i)} \sum_{x\in\mathcal{S}_i} \pi_x^\varepsilon \tau_x^{\varepsilon,j} = \frac{\left\langle \boldsymbol{\delta}^i, \boldsymbol{\tau}^{\varepsilon,j} \right\rangle_{\boldsymbol{\pi}^\varepsilon}}{\boldsymbol{\pi}^\varepsilon(\mathcal{S}_i)}. \tag{5.0.2}$$

Recall the notation of the mean first transition time between two states $x, y \in \mathcal{S}$,

$$m^{\varepsilon}_{x,y} \coloneqq \mathbb{E}_x \inf\left\{k \geq 0 |\ Z^{\varepsilon}_k = y\right\}.$$

To prove the main theorem of this chapter we must assume an additional condition for the Markov chain Z^{ε}:

(B) Assume that there is an $1 \leq i \leq n$ such that for all $y \in \mathcal{S}$

$$\limsup_{\varepsilon \to 0} \sum_{z \in \mathcal{S} \setminus \{\mathfrak{m}_i, y\}} \frac{p^{\varepsilon}_{\mathfrak{m}_i, z}}{1 - p^{\varepsilon}_{\mathfrak{m}_i, \mathfrak{m}_i}} \frac{m^{\varepsilon}_{z,y}}{m^{\varepsilon}_{\mathfrak{m}_i, y}} < 1.$$

We will make further comments on that condition in Section 5.3 where we will also see that the term on the left hand side of this inequality is always less or equal 1.
Now let us state the main theorem. For this recall the definition of $c(\alpha, h)$ in (3.0.2), $c(\alpha, h) = \frac{\alpha}{2c_1(\alpha)h}$, and set $\overline{\mathbf{G}}^{\varepsilon} \coloneqq \frac{\varepsilon^{\alpha}}{c(\alpha,h)}\mathbf{G}$.

Theorem 5.0.1. *Let $\mu^{\varepsilon}_1, \ldots, \mu^{\varepsilon}_n$ and $\lambda^{\varepsilon}_1, \ldots, \lambda^{\varepsilon}_n$ be the eigenvalues of $\overline{\mathbf{G}}^{\varepsilon}$ and $\mathbf{Q}^{\varepsilon}$, respectively, and assume that the condition (B) holds. Then*

$$\mu^{\varepsilon}_1 = \lambda^{\varepsilon}_1 = 0 \quad \textit{and} \quad \lim_{\varepsilon \to 0} \left|(\lambda^{\varepsilon}_i)^{-1} - \mu^{\varepsilon}_i\right| = 0, \quad 2 \leq i \leq n.$$

As opposed to the proof of the continuous analog of that theorem given in [KoMa] it is not possible for us to work directly with the solutions to (5.0.1). We will elaborate on that matter in more detail in Section 5.1. Fortunately, it turns out that these solutions are closely related to the columns of the so called fundamental matrix and deviation matrix which gives us an alternative approach to the derivation of Theorem 5.0.1. Therefore, we start this chapter with the introduction of these matrices. The second part then establishes the connection to the solutions of (5.0.1). Also, we will simultaneously prove results regarding eigenvalues of these quantities.
In the last part we will show how to use these matrices in combination with the results of Chapter 3 regarding mean first transition time to finish the proof of Theorem 5.0.1.

5.1 The Fundamental Matrix and Deviation Matrix for Regular Markov Chains

The aim of this section is to introduce the basic idea behind the fundamental matrix and the deviation matrix for a regular Markov Chain. We will do that simultaneously for discrete- and continuous-time Markov chains with finite state space $\mathcal{S}_d$ and $\mathcal{S}_c$, respectively. This introduction is based on the book [KeSn1] as well as the papers [KeSn2] and [CoDo] where the interested reader can also find the proofs for the results given below.
Let $\boldsymbol{\pi}^d = (\pi^d_x)_{x\in\mathcal{S}_d}$ and $\boldsymbol{\pi}^c = (\pi^c_x)_{x\in\mathcal{S}_c}$ denote the stationary distributions of the these Markov chains and define the matrices $\mathbf{\Pi}^d$ and $\mathbf{\Pi}^c$ by

$$(\mathbf{\Pi}^d)_{x,y} \coloneqq \pi^d_y,\ x,y\in\mathcal{S}_d \quad \text{and} \quad (\mathbf{\Pi}^c)_{x,y} \coloneqq \pi^c_y,\ x,y\in\mathcal{S}_c,$$

i.e. each row of these matrices is the stationary distribution for the respective chain. We now state the central definition.

Definition 5.1.

(i) The *fundamental matrix* $\mathbf{F}^d$ of a regular discrete-time Markov chain $X = (X_k)_{k\geq 0}$ with the transition matrix $\mathbf{P}$ is defined by

$$\mathbf{F}^d \coloneqq (\mathbf{I} - (\mathbf{P} - \mathbf{\Pi}^d))^{-1}.$$

(ii) The *fundamental matrix* $\mathbf{F}^c$ of a regular continuous-time Markov chain $Y = (Y_t)_{t\geq 0}$ with the generator matrix $\mathbf{Q}$ is defined by

$$\mathbf{F}^c \coloneqq (\mathbf{\Pi}^c - \mathbf{Q})^{-1}.$$

These matrices also possess the following representations:

$$\mathbf{F}^d = \mathbf{\Pi}^d + \sum_{k=0}^{\infty}\left(\mathbf{P}^{(k)} - \mathbf{\Pi}^d\right) \quad \text{and} \quad \mathbf{F}^c = \mathbf{\Pi}^c + \int_0^\infty (\mathbf{P}(t) - \mathbf{\Pi}^c)\,dt, \tag{5.1.1}$$

where $(\mathbf{P}(t))_{t\geq 0}$ is the family of transition matrices of Y. As it is pointed out in [CoDo] some authors (see e.g. [Sys], [Whi]) use the terminology "fundamental matrix" for the matrices $\sum_{k=0}^{\infty}\left(\mathbf{P}^{(k)} - \mathbf{\Pi}^d\right)$ and $\int_0^\infty (\mathbf{P}(t) - \mathbf{\Pi}^c)\,dt$, respectively. Clearly, the formulas

(5.1.1) show that the columns of these matrices differ from the ones of the fundamental matrix as we defined it only by a shift with a vector of the form $a\mathbf{1}$, i.e. with a vector from the kernel of the generator. Therefore, there is not much of a difference between these definitions. We will use the same terminology as in [CoDo] and state the definition of the deviation matrix now.

Definition 5.2. Let $\mathbf{F}$ denote the fundamental matrix of a (discrete- or continuous-time) Markov chain X and let $\mathbf{\Pi}$ the matrix where each row equals the stationary distribution. Then

$$\mathbf{D} := \mathbf{F} - \mathbf{\Pi}$$

is called the *deviation matrix* of X.

Remark: The authors in [KeSn1] gave the following interpretation for the entries $f_{i,j}$ of the fundamental matrix $\mathbf{F}$: If $v_j^{(k)}$ denotes the number of times that the Markov chain visits the state j in the first k steps then the mean $\mathbb{E}_i v_j^{(k)}$ for a given state i differs from the value $k\pi_j$ by approximately $f_{i,j} - \pi_j$ (as $k \to \infty$).

In the next proposition we collect some useful formulas for the fundamental matrix.

Proposition 5.1.1. *Let $\mathbf{F}^d$ and $\mathbf{F}^c$ denote the fundamental matrices of a discrete-time Markov chain with transition matrix $\mathbf{P}$ and a continuous-time Markov chain with generator $\mathbf{Q}$, respectively, and let $\mathbf{\Pi}^d$ and $\mathbf{\Pi}^c$ be given as above. Then:*

(i) $\mathbf{F}^d\mathbf{1} = \mathbf{1}$ *and* $\mathbf{F}^c\mathbf{1} = \mathbf{1}$.

(ii) $(\mathbf{P} - \mathbf{I})\mathbf{F}^d = \mathbf{\Pi}^d - \mathbf{I}$ *and* $\mathbf{Q}\mathbf{F}^c = \mathbf{\Pi}^c - \mathbf{I}$.

Obviously, using Definition 5.2 one can rewrite this proposition in terms of the deviation matrix.

Proposition 5.1.2. *Let $\mathbf{D}^d$ and $\mathbf{D}^c$ denote the deviation matrices of a discrete-time Markov chain with transition matrix $\mathbf{P}$ and a continuous-time Markov chain with generator $\mathbf{Q}$, respectively, and let $\mathbf{\Pi}^d$ and $\mathbf{\Pi}^c$ be given as above. Then:*

(i) $\mathbf{D}^d\mathbf{1} = \mathbf{0}$ *and* $\mathbf{D}^c\mathbf{1} = \mathbf{0}$.

(ii) $(\mathbf{P} - \mathbf{I})\mathbf{D}^d = \mathbf{\Pi}^d - \mathbf{I}$ *and* $\mathbf{Q}\mathbf{D}^c = \mathbf{\Pi}^c - \mathbf{I}$.

Now let us briefly comment on the connection between the matrix of expected first passage times and the fundamental matrix as well as the deviation matrix. Let $\mathcal{S}$ be a finite state space and let τ_y be the first passage time of the state y. Define the matrix $\mathbf{M} = (m_{x,y})_{x,y\in\mathcal{S}}$ by

$$m_{x,y} = \mathbb{E}_x\tau_y.$$

Recall from Section 1.1 the notations $\mathbf{E}$ as well as $\mathbf{A}_{\mathrm{dg}}$ when $\mathbf{A}$ is a given matrix. Generally it can be said that the matrices $\mathbf{F}, \mathbf{P}$ and $\mathbf{M}$ are closely connected to each other. Indeed, assume that the stationary distribution is known, then in [KeSn1], Theorem 4.4.7, the authors state the formula

$$\mathbf{M} = (\mathbf{I} - \mathbf{F} + \mathbf{E}\mathbf{F}_{\mathrm{dg}})(\mathbf{\Pi}_{\mathrm{dg}})^{-1}.$$

On the other hand (see Theorem 4.4.12) they show that if $\mathbf{M}$ is given, then the transition matrix $\mathbf{P}$ can be computed from the identity

$$\mathbf{P} = \mathbf{I} + ((\mathbf{\Pi}_{\mathrm{dg}})^{-1} - \mathbf{E})(\mathbf{M} - \mathbf{M}_{\mathrm{dg}})^{-1}.$$

Our aim is to apply this theory to the Markov chain Z^ε and since we already know quite a bit about expected first transition times from Chapter 3 the next proposition shows how to obtain the deviation matrix from $\mathbf{M}$ and $\mathbf{\Pi}$.

Proposition 5.1.3. *Let X be a (discrete or continuous time) Markov chain and let* $\mathbf{M}$ *and* $\mathbf{\Pi}$ *be defined as above. Then*

$$\mathbf{D} = (\mathbf{\Pi} - \mathbf{I})\mathbf{M}\mathbf{\Pi}_{dg}.$$

For more formulas and further details on that matter see again [KeSn1], [KeSn2] and [CoDo].

We conclude this chapter by showing that, when it comes to eigenvalues, it essentially does not matter if we work with the fundamental matrix or with the deviation matrix.

Proposition 5.1.4. *$\lambda^{\mathbf{F}} = 1$ and $\lambda^{\mathbf{D}} = 0$ are eigenvalues of $\mathbf{F}$ and $\mathbf{D}$, respectively. Moreover, for every eigenvalue $\lambda^{\mathbf{F}}$ of $\mathbf{F}$ such that a corresponding eigenvector $\mathbf{x}^{\mathbf{F}}$ does not belong to $\{a\mathbf{1}|\ a \in \mathbb{C}\}$ one can find an eigenvalue $\lambda^{\mathbf{D}}$ of $\mathbf{D}$ such that*

$$\lambda^{\mathbf{F}} = \lambda^{\mathbf{D}}.$$

Proof. The first statement follows from Proposition 5.1.1 (i) and Proposition 5.1.2 (i). Now, let $\lambda^{\mathbf{F}}$ be an eigenvalue of $\mathbf{F}$ with an eigenvector $\mathbf{x}^{\mathbf{F}} \notin \{a\mathbf{1}|\ a \in \mathbb{C}\}$. Then the vector $\mathbf{y} \coloneqq \mathbf{x}^{\mathbf{F}} - \frac{\langle \boldsymbol{\pi}, \mathbf{x}^{\mathbf{F}} \rangle}{\lambda^{\mathbf{F}}} \mathbf{1}$ is well defined ($\mathbf{F}$ is by definition an inverse matrix and therefore $\lambda^{\mathbf{F}} \neq 0$) and not of the form $a\mathbf{1}$ (for some $a \in \mathbb{C}$) as well. Finally, an easy calculation shows that $\mathbf{y}$ is the eigenvector of $\mathbf{D}$ with the eigenvalue $\lambda^{\mathbf{F}}$.

□

Remark: Although the algebraic multiplicity of the eigenvalue 0 of $\mathbf{D}$, i.e. the multiplicity of these numbers as zeros of the characteristic polynomial, might be larger than 1, we can conclude with the help of Proposition 5.1.2 (ii) that the geometric multiplicity, i.e. the dimension of the eigenspace, is always 1. Indeed, if we assume that there is another vector $\tilde{\mathbf{x}} \neq a\mathbf{1}$ (for some $a \in \mathbb{C}$) such that $\mathbf{D}\tilde{\mathbf{x}} = \mathbf{0}$, then it would follow that

$$\mathbf{\Pi}\tilde{\mathbf{x}} = \tilde{\mathbf{x}}.$$

But $\mathbf{\Pi}$ has precisely 2 eigenvalues, namely 0 with algebraic multiplicity $K - 1$ and 1 with algebraic multiplicity 1. Also, the equation $\mathbf{\Pi}\mathbf{1} = \mathbf{1}$ holds. That implies $\tilde{\mathbf{x}} \in \{a\mathbf{1}|\ a \in \mathbb{C}\}$ which contradicts the assumption.

5.2 The Discrete Poisson Problem, Deviation Matrix and Eigenvalues

5.2.1 Continuous-time Markov Chain Y

Denote by $\boldsymbol{\pi}^{\boldsymbol{Q}}$ the stationary distribution of Y and let $\boldsymbol{\delta}^j = (\delta_i^j)_{i=1}^n$ $(1 \leq j \leq n)$ be j-th unit vector, i.e. $\delta_i^j = 1$ for $i = j$ and 0 otherwise. Then, the analog to (5.0.1) for this case

is the system

$$\mathbf{Q}\boldsymbol{\tau}^j = \pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j \tag{5.2.1}$$

with the unknown vector $\boldsymbol{\tau}^j$. We should start by showing the existence of a solution to this system.

Lemma 5.2.1. *The linear system (5.2.1) possesses a solution.*

Actually, this statement already follows from Proposition 5.1.1 (i) which states that the columns of the fundamental matrix are a solution to this system. But since we will need an explicit formula for the solution we will give a more elementary proof. Also, one should keep in mind that the fundamental matrix is defined as an inverse of a certain matrix of which the existence must be proven as well.

Proof. Let $E_1 = \{a\mathbf{1} |\ a \in \mathbb{C}\}$ and $\boldsymbol{\pi}^{\mathbf{Q}}$ denote the stationary distribution of Y. We decompose $\mathbb{C}^n$ in the direct sum of E_1 and $E_1^\perp$, $\mathbb{C}^n = E_1 \oplus E_1^\perp$, where $E_1^\perp$ is the orthogonal complement of E_1 with respect to the weighted inner product $\langle\cdot,\cdot\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}}$. First, let us show that the restriction $\mathbf{Q}^\perp := \mathbf{Q}|_{E_1^\perp}$ of $\mathbf{Q}$ on $E_1^\perp$ maps $E_1^\perp$ into $E_1^\perp$ and is bijective.
Indeed, using $\sum_{i=1}^n q_{i,j}\pi_i^{\mathbf{Q}} = 0$ we find for arbitrary $\mathbf{x} \in \mathbb{C}^n$

$$\langle \mathbf{Q}\mathbf{x}, \mathbf{1}\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} = \sum_{i=1}^n \left(\sum_{j=1}^n q_{i,j}x_j\right)\pi_i^{\mathbf{Q}} = \sum_{j=1}^n x_j \sum_{i=1}^n q_{i,j}\pi_i^{\mathbf{Q}} = 0$$

and therefore $\mathbf{Q}^\perp(E_1^\perp) \subseteq \mathbf{Q}(\mathbb{C}^n) \subseteq E_1^\perp$.
Moreover, since every $\mathbf{x} \in \mathbb{C}^n$ with $\mathbf{Q}\mathbf{x} = 0$ is an element of E_1 we can deduce $\ker(\mathbf{Q}^\perp) = \{0\}$ and hence $\mathbf{Q}^\perp$ is injective. We also know that by definition one has $E_1 = \ker(\mathbf{Q})$ and that $\mathbb{C}^n = \ker(\mathbf{Q}) \oplus \operatorname{image}(\mathbf{Q}) = E_1 \oplus E_1^\perp$ which implies $E_1^\perp = \operatorname{image}(\mathbf{Q})$. It follows that for every $\mathbf{x} \in E_1^\perp$ we can find $\mathbf{y} = \mathbf{y}_1 + \mathbf{y}_2, \mathbf{y}_1 \in E_1, \mathbf{y}_2 \in E_1^\perp$ such that $\mathbf{x} = \mathbf{Q}\mathbf{y} = \mathbf{Q}\mathbf{y}_2$ and hence $E_1^\perp \subseteq \mathbf{Q}(E_1^\perp)$ which implies the surjectivity.
Now let us show that $\pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j \in E_1^\perp,\ 1 \le j \le n$. For this consider the projection $\mathbf{P}_{E_1} : \mathbb{C}^n \to E_1$ onto the subspace E_1. By definition of E_1 every $\mathbf{x} \in \mathbb{C}^n$ has a representation $\mathbf{x} = a\mathbf{1} + (\mathbf{x} - a\mathbf{1})$ where $a \in \mathbb{C}$ must be chosen in a way such that $(\mathbf{x} - a\mathbf{1}) \in E_1^\perp$. This condition leads to the equation

$$0 = \langle \mathbf{x} - a\mathbf{1}, \mathbf{1}\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} = \sum_{i=1}^n x_i\pi_i^{\mathbf{Q}} - a\sum_{i=1}^n \pi_i^{\mathbf{Q}} = \sum_{i=1}^n x_i\pi_i^{\mathbf{Q}} - a$$

and therefore the projection $\mathbf{P}_{E_1}$ is given by $\mathbf{P}_{E_1}\mathbf{x} = \langle \mathbf{x}, \mathbf{1}\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} \mathbf{1}$. If we denote by $\mathbf{P}_{E_1^\perp}$ the projection onto $E_1^\perp$, then we have $\mathbf{I}_n = \mathbf{P}_{E_1} + \mathbf{P}_{E_1^\perp}$ and

$$\mathbf{P}_{E_1^\perp}(-\boldsymbol{\delta}^j) = (\mathbf{I}_n - \mathbf{P}_{E_1})(-\boldsymbol{\delta}^j) = \left\langle \boldsymbol{\delta}^j, \mathbf{1}\right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} \mathbf{1} - \boldsymbol{\delta}^j = \pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j \tag{5.2.2}$$

and that implies $\pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j \in E_1^\perp$.
That allows us to apply $(\mathbf{Q}^\perp)^{-1}$ to this vector. So set $\boldsymbol{\tau}^j \coloneqq (\mathbf{Q}^\perp)^{-1}(\pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j)$ and since $\mathbf{Q} = \mathbf{Q}^\perp$ on $E_1^\perp$ we find

$$\mathbf{Q}\boldsymbol{\tau}^j = \mathbf{Q}^\perp(\mathbf{Q}^\perp)^{-1}(\pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j) = \pi_j^{\mathbf{Q}}\mathbf{1} - \boldsymbol{\delta}^j \tag{5.2.3}$$

and hence this $\boldsymbol{\tau}^j$ is a solution to (5.2.1).

□

Now denote by $\mathbf{G} = (g_{i,j})_{i,j=1}^n$ the matrix given by the relation

$$g_{i,j} \coloneqq \frac{\left\langle \boldsymbol{\delta}^i, \boldsymbol{\tau}^j\right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}}}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}^{\mathbf{Q}}}} = \tau_i^j. \tag{5.2.4}$$

We can make the following statement regarding the eigenvalues of $\mathbf{G}$. Recall that we assume that all eigenvalues of $\mathbf{Q}$ are real and simple, see condition **(E)** at the end of Section 4.2.

Proposition 5.2.2. *Let $\mu_1, \dots, \mu_n$ be the eigenvalues of $\mathbf{G}$ and $\lambda_1^{\mathbf{Q}}, \dots, \lambda_n^{\mathbf{Q}}$ be the eigenvalues of $\mathbf{Q}$. Then*

$$\mu_1 = 0 \quad \textit{and} \quad \mu_i = \frac{1}{\lambda_i^{\mathbf{Q}}},\ 2 \le i \le n.$$

Proof. As in the proof of Lemma 5.2.1 we denote by $(\mathbf{Q}^\perp)^{-1}$ the inverse of $\mathbf{Q}$ on the space $E_1^\perp$ and by $\mathbf{P}_{E_1^\perp}$ the projection on this space. From the equations (5.2.2) and (5.2.3) it

follows that $\tau^j = -(\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp}\boldsymbol{\delta}^j$. The fact that $\mu_1 = 0$ is given by the relation

$$\begin{aligned}\sum_{j=1}^n g_{i,j} &= \sum_{j=1}^n \frac{\left\langle \boldsymbol{\delta}^i, \boldsymbol{\tau}^j \right\rangle_{\boldsymbol{\pi}\mathbf{Q}}}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}\mathbf{Q}}} \\ &= -\frac{1}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}\mathbf{Q}}} \sum_{j=1}^n \left\langle \boldsymbol{\delta}^i, (\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp}\boldsymbol{\delta}^j \right\rangle_{\boldsymbol{\pi}\mathbf{Q}} \\ &= -\frac{1}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}\mathbf{Q}}} \left\langle \boldsymbol{\delta}^i, (\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp} \sum_{j=1}^n \boldsymbol{\delta}^j \right\rangle_{\boldsymbol{\pi}\mathbf{Q}} \\ &= -\frac{1}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}\mathbf{Q}}} \left\langle \boldsymbol{\delta}^i, (\mathbf{Q}^\perp)^{-1} \underbrace{\mathbf{P}_{E_1^\perp}\mathbf{1}}_{=\mathbf{0}} \right\rangle_{\boldsymbol{\pi}\mathbf{Q}} \\ &= 0 \end{aligned} \tag{5.2.5}$$

for every $1 \leq i \leq n$ which implies $\mathbf{G1} = \mathbf{0}$.

Set $\hat{\boldsymbol{\delta}}^i \coloneqq \frac{\boldsymbol{\delta}^i}{\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}\mathbf{Q}}}$. Then

$$g_{i,j} = \frac{\left\langle \boldsymbol{\delta}^i, \boldsymbol{\tau}^j \right\rangle_{\boldsymbol{\pi}\mathbf{Q}}}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}\mathbf{Q}}} = -\frac{\|\boldsymbol{\delta}^j\|_{\boldsymbol{\pi}\mathbf{Q}}}{\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}\mathbf{Q}}} \left\langle \hat{\boldsymbol{\delta}}^i, (\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp}\hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}\mathbf{Q}}.$$

Now let μ be an arbitrary eigenvalue of $\mathbf{G}$ with corresponding (suitable normalized) eigenvector $\boldsymbol{\phi}$. Define a matrix $\mathbf{H} = (h_{i,j})_{i,j=1}^n$ by $h_{i,j} \coloneqq \left\langle \hat{\boldsymbol{\delta}}^i, (\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp}\hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}\mathbf{Q}}$ and a vector $\boldsymbol{\zeta} = (\zeta_i)_{i=1}^n$ by $\zeta_i \coloneqq -\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}\mathbf{Q}}\phi_i$. Then

$$(\mathbf{H}\boldsymbol{\zeta})_i = \sum_{j=1}^n h_{i,j}\zeta_j = -\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}\mathbf{Q}} \sum_{j=1}^n g_{i,j}\frac{1}{\|\boldsymbol{\delta}^j\|_{\boldsymbol{\pi}\mathbf{Q}}}\zeta_j = -\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}\mathbf{Q}}(\mathbf{G}\boldsymbol{\phi})_i = \mu\boldsymbol{\zeta_i}$$

and hence $\mathbf{H}$ and $\mathbf{G}$ have the same spectrum.

Define $\boldsymbol{\Lambda}^{-1} \coloneqq \operatorname{diag}(0, (\lambda_2^{\mathbf{Q}})^{-1}, \ldots, (\lambda_n^{\mathbf{Q}})^{-1})$. Our goal is to show that $\mathbf{H}$ is similar to $\boldsymbol{\Lambda}^{-1}$. Indeed, let $\boldsymbol{\psi}_j^{\mathbf{Q}}$ denote the eigenvector corresponding $\lambda_j^{\mathbf{Q}}$ normalized such that $(\boldsymbol{\psi}_j^{\mathbf{Q}})_j = 1$. Then, since we assume that all eigenvalues are simple, we can conclude that these eigenvectors form a basis for $\mathbb{C}^n$ and therefore we can find numbers $c_{i,j}$ such that

$$\hat{\boldsymbol{\delta}}^i = \sum_{j=1}^n c_{i,j}\boldsymbol{\psi}_j^{\mathbf{Q}}.$$

Now collect all these numbers in a matrix $\mathbf{C}$, i.e. $\mathbf{C} = (c_{i,j})_{i,j=1}^n$, and define ,furthermore,

$\tilde{\mathbf{C}} = (\tilde{c}_{i,j})_{i,j=1}^n$ by $\tilde{c}_{i,j} := \left\langle \boldsymbol{\psi}_i^{\mathbf{Q}}, \hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}}$. Then

$$(\mathbf{C}\tilde{\mathbf{C}})_{i,j} = \sum_{k=1}^n c_{i,k}\tilde{c}_{k,j} = \left\langle \sum_{k=1}^n c_{i,k}\boldsymbol{\psi}_k^{\mathbf{Q}}, \hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} = \left\langle \hat{\boldsymbol{\delta}}^i, \hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

and hence $\tilde{\mathbf{C}} = \mathbf{C}^{-1}$. Finally,

$$\begin{aligned} h_{i,j} &= \left\langle \hat{\boldsymbol{\delta}}^i, (\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp}\hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} = \left\langle \hat{\boldsymbol{\delta}}^i, (\mathbf{Q}^\perp)^{-1}\mathbf{P}_{E_1^\perp}\sum_{k=1}^n c_{j,k}\boldsymbol{\psi}_k^{\mathbf{Q}} \right\rangle_{\boldsymbol{\pi}^{\mathbf{Q}}} = \sum_{k=2}^n c_{j,k}\frac{1}{\lambda_k^{\mathbf{Q}}}\tilde{c}_{k,i} \\ &= (\mathbf{C}\boldsymbol{\Lambda}^{-1}\mathbf{C}^{-1})_{j,i}. \end{aligned}$$

That concludes the proof. □

It is clear that this matrix $\mathbf{G}$ is closely related to the deviation matrix $\mathbf{D}^{\mathbf{Q}}$. Indeed, since the columns of $\mathbf{G}$ are the solutions to (5.2.1) and $\mathbf{D}^{\mathbf{Q}}$ solves the matrix equation $\mathbf{Q}\mathbf{D}^{\mathbf{Q}} = \boldsymbol{\Pi} - \mathbf{I}_n$ (see Proposition 5.1.2 (ii)) we have

$$\mathbf{G} = \mathbf{D}^{\mathbf{Q}} + \mathbf{A} \tag{5.2.6}$$

where the columns $\mathbf{a}_j$, $1 \leq j \leq n$, of $\mathbf{A}$ are from the kernel of $\mathbf{Q}$, i.e. $\mathbf{a}_j = a_j\mathbf{1}$ for some $a_j \in \mathbb{C}$. This has a nice consequence for the spectra of these matrices.

Proposition 5.2.3. *Let $\mu_1, \ldots, \mu_n$ and $\nu_1, \ldots, \nu_n$ denote the eigenvalues of $\mathbf{G}$ and $\mathbf{D}^{\mathbf{Q}}$, respectively. Then*

$$\mu_1 = \nu_1 = 0 \quad \textit{and} \quad \mu_i = \nu_i, \quad 2 \leq i \leq n. \tag{5.2.7}$$

Proof. Denote by $\mathbf{a} = (a_1, \ldots, a_n)^T$ the vector that defines the matrix $\mathbf{A}$ in (5.2.6).
The fact that 0 is the eigenvalue for both $\mathbf{G}$ and $\mathbf{D}^{\mathbf{Q}}$ follows from Proposition 5.1.2 (i) as well as Proposition 5.2.2 or, more precisely, from formula (5.2.5) from which one can also see that in both cases $\mathbf{1}$ is the corresponding eigenvector. That implies

$$0 = \mathbf{G}\mathbf{1} = \mathbf{D}^{\mathbf{Q}}\mathbf{1} + \mathbf{A}\mathbf{1} = \mathbf{A}\mathbf{1}$$

and hence $\langle \mathbf{a}, \mathbf{1} \rangle = 0$.
Now let $\mu \neq 0$ be another eigenvalue of $\mathbf{G}$ with a corresponding eigenvector $\mathbf{u}$. We set

$\mathbf{v} = \mathbf{u} + \gamma\mathbf{1}$ and determine γ such that $\mathbf{v}$ is an eigenvector of $\mathbf{D^Q}$ corresponding to μ. We find

$$\begin{aligned}\mathbf{D^Q v} &= (\mathbf{G} - \mathbf{A})(\mathbf{u} + \gamma\mathbf{1}) = \mu\mathbf{u} - \langle \mathbf{a}, \mathbf{u}\rangle \mathbf{1} + \gamma\mathbf{G1} + \gamma \langle \mathbf{a}, \mathbf{1}\rangle \mathbf{1} \\ &= \mu\mathbf{u} - \langle \mathbf{a}, \mathbf{u}\rangle \mathbf{1} \\ &\stackrel{!}{=} \mu\mathbf{u} + \mu\gamma\mathbf{1}\end{aligned}$$

and thus $\gamma = -\frac{\langle \mathbf{a},\mathbf{u}\rangle}{\mu}$.

□

5.2.2 Discrete-time Markov Chain Z^ε

Consider again the linear system (5.0.1):

$$(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\boldsymbol{\tau}^{\varepsilon,j} = \boldsymbol{\pi}^\varepsilon(\mathcal{S}_j)\mathbf{1} - \boldsymbol{\delta}^j.$$

where $\delta^j_x = 1$ for $x \in \mathcal{S}_j$ and 0 otherwise. The existence of a solution to this system can be proven similarly to the last section.
Note that the statement of Proposition 5.2.2 is essentially the one of Theorem 5.0.1 in terms of the Markov chain Y instead of Z^ε.

Let us now briefly comment on why, in our opinion, it is not possible to use similar methods to prove Theorem 5.0.1 as we did in the last section and to what was done in [KoMa] or [Ko2], Chapter 8, Section 2.
Again, 0 is an eigenvalue of $\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|}$ with the eigenspace $E_1 = \{a\mathbf{1} |\, a \in \mathbb{C}\}$. Let $E_1^\perp$ denote the orthogonal complement of E_1 with respect to the inner product $\langle\cdot,\cdot\rangle_{\boldsymbol{\pi}^\varepsilon}$ and let $(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})^\perp$ and $\mathbf{P}^\perp_{E_1}$ be the restriction of $(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})$ on $E_1^\perp$ and the projection on $E_1^\perp$, respectively. As in (5.2.4) we define $\mathbf{G}^\varepsilon = (g^\varepsilon_{i,j})_{i,j=1}^n$ by

$$g^\varepsilon_{i,j} := \frac{\left\langle \boldsymbol{\delta}^i, \boldsymbol{\tau}^{\varepsilon,j}\right\rangle_{\boldsymbol{\pi}^\varepsilon}}{\|\boldsymbol{\delta}^i\|^2_{\boldsymbol{\pi}^\varepsilon}} = \frac{\left\langle \boldsymbol{\delta}^i, \boldsymbol{\tau}^{\varepsilon,j}\right\rangle_{\boldsymbol{\pi}^\varepsilon}}{\boldsymbol{\pi}^\varepsilon(\mathcal{S}_i)}. \tag{5.2.8}$$

Denote $\hat{\boldsymbol{\delta}}^i := \frac{\boldsymbol{\delta}^i}{\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}^\varepsilon}}$. We then can write

$$g^\varepsilon_{i,j} = -\frac{\|\boldsymbol{\delta}^j\|_{\boldsymbol{\pi}^\varepsilon}}{\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}^\varepsilon}} \left\langle \hat{\boldsymbol{\delta}}^i, ((\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})^\perp)^{-1}\mathbf{P}^\perp_1\hat{\boldsymbol{\delta}}^j\right\rangle_{\boldsymbol{\pi}^\varepsilon}.$$

Now define an auxiliary matrix $\mathbf{H}^\varepsilon = (h^\varepsilon_{i,j})_{i,j=1}^n$ by

$$h^\varepsilon_{i,j} := \left\langle \hat{\boldsymbol{\delta}}^i, ((\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})^\perp)^{-1}\mathbf{P}_1^\perp \hat{\boldsymbol{\delta}}^j \right\rangle_{\boldsymbol{\pi}^\varepsilon}.$$

For every eigenvalue μ^ε of $\mathbf{G}^\varepsilon$ with an eigenvector $\mathbf{x}^\varepsilon$ let us construct a new vector $\mathbf{y}^\varepsilon$ by $y^\varepsilon_i := -\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}^\varepsilon} x^\varepsilon_i$. Then

$$(\mathbf{H}^\varepsilon \mathbf{y}^\varepsilon)_i = \sum_{j=1}^n h^\varepsilon_{i,j} y^\varepsilon_j = -\sum_{j=1}^n \frac{\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}^\varepsilon}}{\|\boldsymbol{\delta}^j\|_{\boldsymbol{\pi}^\varepsilon}} g^\varepsilon_{i,j} y^\varepsilon_j = -\|\boldsymbol{\delta}^i\|_{\boldsymbol{\pi}^\varepsilon} \sum_{j=1}^n g^\varepsilon_{i,j} x^\varepsilon_j = \mu^\varepsilon y^\varepsilon_i.$$

Hence, the spectra of $\mathbf{G}^\varepsilon$ and $\mathbf{H}^\varepsilon$ coincide. Also, as in (5.2.5) one can see that 0 is an eigenvalue of $\mathbf{G}^\varepsilon$ (and therefore of $\mathbf{H}^\varepsilon$).

Now the goal would be to show that the eigenvalues μ^ε_i of $\frac{\varepsilon^\alpha}{c(\alpha,h)}\mathbf{H}^\varepsilon$ approximate, as $\varepsilon \to 0$, the values $(\lambda^\varepsilon_i)^{-1}$ where λ^ε_i is one of the eigenvalues from Theorem 4.2.2 (i). If we would follow the ideas presented in the proof of Theorem 1.2 in [KoMa], Chapter 3, we would end up with an equation of the form

$$h^\varepsilon_{i,j} = \left(\mathbf{C}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{C}\right)_{i,j} + r^\varepsilon_{i,j}$$

for a certain matrix $\mathbf{C}$ and $\boldsymbol{\Lambda}^{-1} = \operatorname{diag}(0, (\lambda^\varepsilon_2)^{-1}, \dots, (\lambda^\varepsilon_n)^{-1})$. However, in contrast to [KoMa] we are not able to show that the remainder term $r^\varepsilon_{i,j}$ vanishes in the limit of small ε.

Let us introduce an alternative approach. We define $|\mathcal{S}| \times n$-rectangular matrices $\mathbf{T}^\varepsilon := (\boldsymbol{\tau}^{\varepsilon,1} \cdots \boldsymbol{\tau}^{\varepsilon,n})$, $\tilde{\boldsymbol{\Pi}}^\varepsilon := \left(\sum_{y\in\mathcal{S}_j} \pi^\varepsilon_y\right)_{x\in\mathcal{S}}^{1\le j\le n}$ and $\boldsymbol{\Delta} = (\boldsymbol{\delta}^1 \cdots \boldsymbol{\delta}^n)$. Then (5.0.1) can be written as the matrix equation

$$(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\mathbf{T}^\varepsilon = \tilde{\boldsymbol{\Pi}}^\varepsilon - \boldsymbol{\Delta}.$$

Similarly to the last section, a comparison with Proposition 5.1.2 indicates a connection between the deviation matrix $\mathbf{D}^\varepsilon$ of Z^ε and the solutions of the system (5.0.1). Indeed, if we define $\tilde{\mathbf{T}}^\varepsilon := \mathbf{D}^\varepsilon\boldsymbol{\Delta}$ and denote by $\boldsymbol{\Pi}^\varepsilon$ the $|\mathcal{S}| \times |\mathcal{S}|$-matrix where each row consists of the stationary distribution $\boldsymbol{\pi}^\varepsilon$, i.e. $(\boldsymbol{\Pi}^\varepsilon)_{x,y} = \pi^\varepsilon_y$ for all $x \in \mathcal{S}$. Then $\tilde{\boldsymbol{\Pi}}^\varepsilon = \boldsymbol{\Pi}^\varepsilon\boldsymbol{\Delta}$ and with Proposition 5.1.2 (ii) we find

$$(\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\tilde{\mathbf{T}}^\varepsilon = (\mathbf{P}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\mathbf{D}^\varepsilon\boldsymbol{\Delta} = (\boldsymbol{\Pi}^\varepsilon - \mathbf{I}_{|\mathcal{S}|})\boldsymbol{\Delta} = \tilde{\boldsymbol{\Pi}}^\varepsilon - \boldsymbol{\Delta},$$

i.e. if we denote by $\tilde{\boldsymbol{\tau}}^{\varepsilon,i}$ the columns of $\tilde{\mathbf{T}}^{\varepsilon}$, then these vectors are solutions to (5.0.1). From the identity $\ker(\mathbf{P}^{\varepsilon} - \mathbf{I}_{|\mathcal{S}|}) = \{a\mathbf{1} |\ a \in \mathbb{C}\}$ it follows that the vectors $\boldsymbol{\tau}^{\varepsilon,i}$ and $\tilde{\boldsymbol{\tau}}^{\varepsilon,i}$ are connected by the equation

$$\boldsymbol{\tau}^{\varepsilon,i} = \tilde{\boldsymbol{\tau}}^{\varepsilon,i} + a_i^{\varepsilon}\mathbf{1} \quad \text{for some } a_i^{\varepsilon} \in \mathbb{C}. \tag{5.2.9}$$

Analogously to (5.2.8) let us consider $\tilde{\mathbf{G}}^{\varepsilon} = (\tilde{g}_{i,j}^{\varepsilon})_{i,j=1}^{n}$ given by

$$\tilde{g}_{i,j}^{\varepsilon} \coloneqq \frac{\left\langle \boldsymbol{\delta}^i, \tilde{\boldsymbol{\tau}}^{\varepsilon,j} \right\rangle_{\boldsymbol{\pi}^{\varepsilon}}}{\pi^{\varepsilon}(\mathcal{S}_j)}. \tag{5.2.10}$$

A straightforward calculation yields

$$\mathbf{G}^{\varepsilon} = \tilde{\mathbf{G}}^{\varepsilon} + \mathbf{A}^{\varepsilon} \quad \text{where } \mathbf{A}^{\varepsilon} = (a_j^{\varepsilon})_{i,j=1}^{n} \quad \text{and the } a_j^{\varepsilon} \text{ are uniquely defined in (5.2.9).}$$

We conclude by showing that, not surprisingly, the spectrum of $\mathbf{G}^{\varepsilon}$ essentially coincides with the spectrum of $\tilde{\mathbf{G}}^{\varepsilon}$. That is the analog to Proposition 5.2.3 only that in this case we do not know a priori if $\left\langle (a_1^{\varepsilon}, \dots, a_n^{\varepsilon})^T, \mathbf{1} \right\rangle$ is already an eigenvalue. If that is the case then the spectrum changes only in regard to the first eigenvalue.

Proposition 5.2.4. *Denote* $\mathbf{a}^{\varepsilon} = (a_1^{\varepsilon}, \dots, a_n^{\varepsilon})^T$ *and let* $\sigma(\mathbf{G}^{\varepsilon}) = \{0, \mu_2^{\varepsilon}, \dots, \mu_n^{\varepsilon}\}$ *be the spectrum of* $\mathbf{G}^{\varepsilon}$.

(i) If $\langle \mathbf{a}^{\varepsilon}, \mathbf{1} \rangle \in \sigma(\mathbf{G}^{\varepsilon})$, *then* $\sigma(\tilde{\mathbf{G}}^{\varepsilon}) = \sigma(\mathbf{G}^{\varepsilon})$.

(ii) If $\langle \mathbf{a}^{\varepsilon}, \mathbf{1} \rangle \notin \sigma(\mathbf{G}^{\varepsilon})$, *then* $\sigma(\tilde{\mathbf{G}}^{\varepsilon}) = \{-\langle \mathbf{a}^{\varepsilon}, \mathbf{1} \rangle, \mu_2^{\varepsilon}, \dots, \mu_n^{\varepsilon}\}$.

Proof. To prove the statement (i) one just has to repeat the arguments from the proof of Proposition 5.2.3. So let us prove (ii). Assume $\mathbf{G}^{\varepsilon}$ possesses n eigenvalues $\mu_1^{\varepsilon} = 0, \mu_2^{\varepsilon}, \dots, \mu_n^{\varepsilon}$ with $\mu_k^{\varepsilon} \neq \langle \mathbf{a}^{\varepsilon}, \mathbf{1} \rangle = \sum_{j=1}^{n} a_j^{\varepsilon}$ for all $1 \leq k \leq n$. We already know that $\mathbf{1}$ is an eigenvector of $\mathbf{G}^{\varepsilon}$ corresponding to the eigenvalue 0. Using $\mathbf{G}^{\varepsilon} = \tilde{\mathbf{G}}^{\varepsilon} + \mathbf{A}^{\varepsilon}$ then yields

$$\tilde{\mathbf{G}}^{\varepsilon}\mathbf{1} = (\mathbf{G}^{\varepsilon} - \mathbf{A}^{\varepsilon})\mathbf{1} = -\langle \mathbf{a}^{\varepsilon}, \mathbf{1} \rangle \mathbf{1}.$$

Hence, $\mathbf{1}$ is an eigenvector of $\tilde{\mathbf{G}}^{\varepsilon}$ corresponding to the eigenvalue $-\langle \mathbf{a}^{\varepsilon}, \mathbf{1} \rangle$.

□

5.3 The Deviation Matrix and Mean Transition Times

As we have seen so far, deviation matrices and solutions to certain inhomogeneous linear systems of equations are closely connected. Moreover, the authors in [KoMa] mention that the solutions to these systems of equations are related to mean life times of the process X^ε in the domains of attraction. This is supported by the formula given in Proposition 5.1.3 that connects the deviation matrix with the matrix $\mathbf{M}$ of mean transition times. For the Markov chain Y with generator $\mathbf{Q}$, state space $\mathcal{M}$ and stationary distribution $\boldsymbol{\pi}^{\mathbf{Q}} = (\pi_x^{\mathbf{Q}})_{x\in\mathcal{M}}$ this is simply

$$\mathbf{D}^{\mathbf{Q}} = (\boldsymbol{\Pi}^{\mathbf{Q}} - \mathbf{I}_n)\mathbf{M}^{\mathbf{Q}}\boldsymbol{\Pi}^{\mathbf{Q}}_{\mathrm{dg}}.$$

Computing the entries of $\mathbf{D}^{\mathbf{Q}}$ with that formula yields

$$d^{\mathbf{Q}}_{x,y} = \pi^{\mathbf{Q}}_y \left(\sum_{z\in\mathcal{M}} \pi^{\mathbf{Q}}_z m^{\mathbf{Q}}_{z,y} - m^{\mathbf{Q}}_{x,y} \right), \quad x,y \in \mathcal{M}. \tag{5.3.1}$$

Note that, because of $m^{\mathbf{Q}}_{x,x} = 0$ for all x, in the case $x = y$ this formula simplifies to

$$d^{\mathbf{Q}}_{y,y} = \pi^{\mathbf{Q}}_y \sum_{z\in\mathcal{M}} \pi^{\mathbf{Q}}_z m^{\mathbf{Q}}_{z,y}.$$

Now let us consider the situation for the Markov chain Z^ε. Recall the definition of the matrix $\tilde{\mathbf{G}}^\varepsilon$ in (5.2.10) which includes the columns of the matrix $\tilde{\mathbf{T}}^\varepsilon = \mathbf{D}^\varepsilon\boldsymbol{\Delta}$ where $\mathbf{D}^\varepsilon$ denotes the deviation matrix of Z^ε. An easy computation yields

$$\tilde{g}^\varepsilon_{i,j} = \frac{\sum_{x\in\mathcal{S}_i} \pi^\varepsilon_x \left(\sum_{y\in\mathcal{S}_j} d^\varepsilon_{x,y} \right)}{\pi^\varepsilon(\mathcal{S}_j)}.$$

Denote $\sigma^\varepsilon_y = \min\{k \geq 0 |\ Z^\varepsilon_k = y\}$ and $m^\varepsilon_{x,y} = \mathbb{E}_x\sigma^\varepsilon_y$. Then, analogously to (5.3.1) we have for $x, y \in \mathcal{S}$

$$d^\varepsilon_{y,y} = \pi^\varepsilon_y \sum_{z\in\mathcal{S}} \pi^\varepsilon_z m^\varepsilon_{z,y} \quad \text{and} \quad d^\varepsilon_{x,y} = \pi^\varepsilon_y \left(\sum_{z\in\mathcal{S}} \pi^\varepsilon_z m^\varepsilon_{z,y} - m^\varepsilon_{x,y} \right).$$

Now we want to connect the matrices $\tilde{\mathbf{G}}^\varepsilon$ and $\mathbf{D}^Q$ in the limit $\varepsilon \to 0$. To do so we first have to rescale the respective quantities for the chain Z^ε with the scaling factor $\frac{\varepsilon^\alpha}{c(\alpha,h)}$ where $c(\alpha, h)$ is defined in (3.0.2). So set

$$\overline{m}^\varepsilon_{x,y} := \frac{\varepsilon^\alpha}{c(\alpha,h)} m^\varepsilon_{x,y}$$

and with that

$$\overline{d}^{\varepsilon}_{y,y} \coloneqq \pi^{\varepsilon}_y \sum_{z\in\mathcal{S}} \pi^{\varepsilon}_z \overline{m}^{\varepsilon}_{z,y} \quad \text{and} \quad \overline{d}^{\varepsilon}_{x,y} \coloneqq \pi^{\varepsilon}_y \left(\sum_{z\in\mathcal{S}} \pi^{\varepsilon}_z \overline{m}^{\varepsilon}_{z,y} - \overline{m}^{\varepsilon}_{x,y} \right) \tag{5.3.2}$$

as well as a matrix $\overline{\widetilde{\mathbf{G}}}^{\varepsilon} = \left(\overline{\widetilde{g}}^{\varepsilon}_{i,j}\right)_{i,j=1}^{n}$ given by

$$\overline{\widetilde{g}}^{\varepsilon}_{i,j} \coloneqq \frac{\sum_{x\in\mathcal{S}_i} \pi^{\varepsilon}_x \left(\sum_{y\in\mathcal{S}_j} \overline{d}^{\varepsilon}_{x,y} \right)}{\pi^{\varepsilon}(\mathcal{S}_j)}. \tag{5.3.3}$$

Recall the assumption (B) from the beginning of this chapter:

(B) Assume that there is an $1 \le i \le n$ such that for all $y \in \mathcal{S}$

$$\limsup_{\varepsilon\to 0} \sum_{z\in\mathcal{S}\setminus\mathfrak{m}_i, y} \frac{p^{\varepsilon}_{\mathfrak{m}_i,z}}{1-p^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_i}} \frac{m^{\varepsilon}_{z,y}}{m^{\varepsilon}_{\mathfrak{m}_i,y}} < 1.$$

The proof of the following lemma shows why we need this condition.

Lemma 5.3.1. *Assume that condition (B) holds. Then there exists a constant $C = C(\alpha, h, \delta)$ such that for all $0 < \varepsilon < \varepsilon_0$ and all $x, y \in \mathcal{S}$*

$$\overline{m}^{\varepsilon}_{x,y} \le C.$$

Proof. Consider the linear system in Proposition A.1 and let us take a look at the equation for a starting point $\mathfrak{m}_i \in \mathcal{M}$ for which condition (B) holds, i.e. for this $\mathfrak{m}_i$ and $y \in \mathcal{S}\setminus\{\mathfrak{m}_i\}$ consider

$$m^{\varepsilon}_{\mathfrak{m}_i,y} = 1 + \sum_{x\in\mathcal{S}\setminus\{y\}} p^{\varepsilon}_{\mathfrak{m}_i,x} m^{\varepsilon}_{x,y}$$

or, equivalently,

$$1 - \sum_{x\in\mathcal{S}\setminus\{y,\mathfrak{m}_i\}} \frac{p^{\varepsilon}_{\mathfrak{m}_i,x}}{1-p^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_i}} \frac{m^{\varepsilon}_{x,y}}{m^{\varepsilon}_{\mathfrak{m}_i,y}} = \frac{1}{m^{\varepsilon}_{\mathfrak{m}_i,y}(1-p^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_i})}. \tag{5.3.4}$$

Now, condition (B) implies that the left hand side of this equation is bounded from below by a constant $c_1 > 0$. On the other hand, Proposition 2.2.1 (iii) shows that there is a constant $c_2 > 0$ such that $c_2 \frac{\varepsilon^{\alpha}}{c(\alpha,h)} \le 1 - p^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_i}$ and thus

$$\overline{m}^{\varepsilon}_{\mathfrak{m}_i,y} \le \frac{1}{c_1 c_2}. \tag{5.3.5}$$

Moreover, equation (5.3.4) can also be written as

$$(1 - p^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_i}) - \sum_{x \in \mathcal{S}\setminus\{y,\mathfrak{m}_i\}} p^{\varepsilon}_{\mathfrak{m}_i,x} \frac{m^{\varepsilon}_{x,y}}{m^{\varepsilon}_{\mathfrak{m}_i,y}} = \frac{1}{m^{\varepsilon}_{\mathfrak{m}_i,y}} \tag{5.3.6}$$

The right hand side is always larger than 0 and therefore the left hand side is dominated by $1 - p^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_i}$ which, as we stated above, is of order $O(\varepsilon^{\alpha})$ and likewise are the terms $p^{\varepsilon}_{\mathfrak{m}_i,x}$ for $x \in \mathcal{S} \setminus \{y, \mathfrak{m}_i\}$. That implies the existence of a constant c_3 such that

$$\frac{m^{\varepsilon}_{x,y}}{m^{\varepsilon}_{\mathfrak{m}_i,y}} \leq c_3.$$

Combining that with (5.3.5) finishes the proof.

□

Remark on condition (B): This condition seems to be very artificial and is only used for the argument in the proof of the previous lemma. In fact, since one has relatively many degrees of freedom in the construction of the Markov chain Z^{ε} we strongly believe that one can always find a combination of discretization parameters h, δ, R and a state space $\mathcal{S}$ in the way we did in Chapter 2 such that condition (B) holds. However, a rigorous proof is still an open problem.

Moreover, one could also try a different approach for a proof of inequality (5.3.5). For example one could try to use Cramer's rule to solve the linear system that determines the vector of mean transition times $\mathbf{m}^{\varepsilon}_y = (m^{\varepsilon}_{x,y})_{x\in\mathcal{S}}$ for a given $y \in \mathcal{S} \setminus \{\mathfrak{m}_i\}$ (see Proposition A.1):

$$m^{\varepsilon}_{x,y} = \begin{cases} 0, & y = x, \\ 1 + \sum_{z \neq y} p^{\varepsilon}_{x,z} m^{\varepsilon}_{z,y}, & y \neq x. \end{cases}$$

The coefficient matrix $\mathbf{B}^{\varepsilon}$ for this system is given by

$$(\mathbf{B}^{\varepsilon})_{x,z} \coloneqq \begin{cases} (\mathbf{I}_{|\mathcal{S}|} - \mathbf{P}^{\varepsilon})_{x,z}, & x, z \neq y, \\ 1, & x = z = y, \\ 0, & \text{otherwise.} \end{cases}$$

If we denote by $\mathbf{B}^{\varepsilon}_{\mathfrak{m}_i}$ the matrix obtained from $\mathbf{B}^{\varepsilon}$ by replacing the column indexed by $\mathfrak{m}_i$ with the vector $\mathbf{b} \coloneqq (1, \dots, 1, 0, 1 \dots, 1)^T$, then $m^{\varepsilon}_{\mathfrak{m}_i,y}$ is given by

$$m^{\varepsilon}_{\mathfrak{m}_i,y} = \frac{\det \mathbf{B}^{\varepsilon}_{\mathfrak{m}_i}}{\det \mathbf{B}^{\varepsilon}}. \tag{5.3.7}$$

The heuristic idea is that every column of the matrix $\mathbf{B}^{\varepsilon}$ except for the one indexed by the state y contains entries of order $O(\varepsilon^{\alpha})$. So if we "delete" one of these columns and replace it with a column of order $O(1)$ then the fraction in (5.3.7) should be of order $O(\varepsilon^{-\alpha})$. More precisely, if we denote by $\mathbf{B}^{\varepsilon}(v,w)$ the $(|\mathcal{S}|-1) \times (|\mathcal{S}|-1)$ matrix that is obtained from $\mathbf{B}^{\varepsilon}$ by erasing the row indexed by v and the column indexed by w (and similar $\mathbf{B}^{\varepsilon}_{\mathfrak{m}_i}(v,w)$), then expanding about the column indexed by the state $\mathfrak{m}_i$, which in the case of $\mathbf{B}^{\varepsilon}_{\mathfrak{m}_i}$ is equal to $\mathbf{b}$, yields

$$\det \mathbf{B}^{\varepsilon}_{\mathfrak{m}_i} = \sum_{v \in \mathcal{S} \setminus \{y\}} s(v) \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i) \quad \text{and} \quad \det \mathbf{B}^{\varepsilon} = \sum_{v \in \mathcal{S} \setminus \{y\}} s(v) (\mathbf{B}^{\varepsilon})_{v,\mathfrak{m}_i} \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i)$$

where $s(v) = \pm 1$ just gives the correct algebraic sign. The coefficients $(\mathbf{B}^{\varepsilon})_{v,\mathfrak{m}_i}$ are either of order $O(\varepsilon^{\alpha})$ in the case $\mathfrak{m}_i \neq y^*(v)$ or of order $O(1)$ in the case $\mathfrak{m}_i = y^*(v)$. Let $\mathcal{S}_y(\mathfrak{m}_i) \coloneqq \{x \in \mathcal{S} \setminus \{y\} \mid \mathfrak{m}_i = y^*(x)\}$. Then there are numbers $\tilde{b}^{\varepsilon}_v, v \in \mathcal{S}$, such that $\tilde{b}^{\varepsilon}_{v,y} = O(1)$ and

$$m^{\varepsilon}_{\mathfrak{m}_i,y} = \frac{1}{\varepsilon^{\alpha}} \frac{\sum_{v \in \mathcal{S}} s(v) \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i)}{\sum_{v \in \mathcal{S}_y(\mathfrak{m}_i)} s(v) \tilde{b}^{\varepsilon}_v \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i) + \sum_{v \notin \mathcal{S}_y(\mathfrak{m}_i)} s(v) \frac{\tilde{b}^{\varepsilon}_v}{\varepsilon^{\alpha}} \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i)}.$$

However, it is still open to show that the factor

$$\frac{\sum_{v \in \mathcal{S}} s(v) \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i)}{\sum_{v \in \mathcal{S}_y(\mathfrak{m}_i)} s(v) \tilde{b}^{\varepsilon}_v \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i) + \sum_{v \notin \mathcal{S}_y(\mathfrak{m}_i)} s(v) \frac{\tilde{b}^{\varepsilon}_v}{\varepsilon^{\alpha}} \det \mathbf{B}^{\varepsilon}(v, \mathfrak{m}_i)}$$

cannot explode as $\varepsilon \to 0$ although, similar to the argumentation regarding condition (B), it always seems possible to choose the parameters in a way such that this fraction is bounded from above.

To finish the proof of Theorem 5.0.1 we need the following propositions.

Proposition 5.3.2. *For all $1 \leq i, j \leq n$ we have*

$$\lim_{\varepsilon \to 0} \overline{\overline{g}}^{\varepsilon}_{i,j} = d^{\mathbf{Q}}_{\mathfrak{m}_i,\mathfrak{m}_j}.$$

Proof. In the first step we insert (5.3.2) into (5.3.3), expand it and then rearrange the terms. We derive

$$\begin{aligned}\overline{\tilde{g}}^{\varepsilon}_{j,j} = \frac{1}{\pi^{\varepsilon}(\mathcal{S}_j)} \Bigg[& \pi^{\varepsilon}_{\mathfrak{m}_j} \left(\pi^{\varepsilon}_{\mathfrak{m}_j} \sum_{i=1}^{n} \pi^{\varepsilon}_{\mathfrak{m}_i} \overline{m}^{\varepsilon}_{\mathfrak{m}_i,\mathfrak{m}_j} \right) \\ & + \pi^{\varepsilon}_{\mathfrak{m}_j} \left(\pi^{\varepsilon}_{\mathfrak{m}_j} \sum_{z \in \mathcal{S}\setminus\mathcal{M}} \pi^{\varepsilon}_{z} \overline{m}^{\varepsilon}_{z,\mathfrak{m}_j} \right) \\ & + \pi^{\varepsilon}_{\mathfrak{m}_j} \left(\sum_{y \in \mathcal{S}_j\setminus\{\mathfrak{m}_j\}} \pi^{\varepsilon}_{y} \left(\sum_{x \in \mathcal{S}} \pi^{\varepsilon}_{x} \overline{m}^{\varepsilon}_{x,y} - \overline{m}^{\varepsilon}_{\mathfrak{m}_j,y} \right) \right) \\ & + \sum_{z \in \mathcal{S}_j\setminus\{\mathfrak{m}_j\}} \pi^{\varepsilon}_{z} \left(\pi^{\varepsilon}_{z} \sum_{x \in \mathcal{S}} \pi^{\varepsilon}_{x} \overline{m}^{\varepsilon}_{x,z} + \sum_{y \in \Omega_j\setminus\{z\}} \pi^{\varepsilon}_{y} \left(\sum_{x \in \mathcal{S}} \pi^{\varepsilon}_{x} \overline{m}^{\varepsilon}_{x,yu} - \overline{m}^{\varepsilon}_{z,y} \right) \right) \Bigg] \end{aligned} \tag{5.3.8}$$

and for $i \neq j$

$$\begin{aligned}\overline{\tilde{g}}^{\varepsilon}_{i,j} = \frac{1}{\pi^{\varepsilon}(\mathcal{S}_i)} \Bigg[& \sum_{x \in \mathcal{S}_i} \pi^{\varepsilon}_{x} \left(\pi^{\varepsilon}_{\mathfrak{m}_j} \sum_{k=1}^{n} \pi^{\varepsilon}_{\mathfrak{m}_k} \overline{m}^{\varepsilon}_{\mathfrak{m}_k,\mathfrak{m}_j} - \pi^{\varepsilon}_{\mathfrak{m}_j} \overline{m}^{\varepsilon}_{x,\mathfrak{m}_j} \right) \\ & + \sum_{x \in \mathcal{S}_i} \pi^{\varepsilon}_{x} \left(\pi^{\varepsilon}_{\mathfrak{m}_j} \sum_{z \in \mathcal{S}\setminus\mathcal{M}} \pi^{\varepsilon}_{z} \overline{m}^{\varepsilon}_{z,\mathfrak{m}_j} + \sum_{y \in \mathcal{S}_j\setminus\{\mathfrak{m}_j\}} \sum_{k=1}^{n} \pi^{\varepsilon}_{y} \pi^{\varepsilon}_{\mathfrak{m}_k} \overline{m}^{\varepsilon}_{\mathfrak{m}_k,y} \right) \\ & + \sum_{x \in \mathcal{S}_i} \pi^{\varepsilon}_{x} \left(\sum_{y \in \mathcal{S}_j\setminus\{\mathfrak{m}_j\}} \sum_{z \in \mathcal{S}\setminus\mathcal{M}} \pi^{\varepsilon}_{y} \pi^{\varepsilon}_{z} \overline{m}^{\varepsilon}_{z,y} \right) \\ & - \sum_{x \in \mathcal{S}_i} \pi^{\varepsilon}_{x} \left(\sum_{y \in \mathcal{S}_j\setminus\{\mathfrak{m}_j\}} \pi^{\varepsilon}_{y} \overline{m}^{\varepsilon}_{x,y} \right) \Bigg]. \end{aligned} \tag{5.3.9}$$

Now, by Proposition 2.3.1 (i) and Lemma 5.3.1 we see that all the summands on the right hand side in (5.3.8) and (5.3.9) except for those in the first line vanish in the limit $\varepsilon \to 0$. Furthermore, Proposition 3.2.3 and again Lemma 5.3.1 show that the expressions in the first line converge to the ones given in (5.3.1).

□

Proposition 5.3.3. *Let $\mu^{\varepsilon}_1, \ldots, \mu^{\varepsilon}_n$ and $\mu^{\mathbf{D}}_1 = 0, \ldots, \mu^{\mathbf{D}}_n$ be the eigenvalues of $\overline{\tilde{\mathbf{G}}}^{\varepsilon}$ and $\mathbf{D}^{\mathbf{Q}}$, respectively, and assume condition (B) holds. Then*

$$\lim_{\varepsilon \to 0} \mu^{\varepsilon}_i = \mu^{\mathbf{D}}_i, \quad 2 \leq i \leq n.$$

Note that, in view of Proposition 5.2.4, we cannot necessarily make a statement about μ_1^ϵ but we can "replace" this eigenvalue with 0.

Proof. Clearly, the assertion of this proposition is a consequence of Proposition 5.3.2 since the eigenvalues depend continuously on the entries of the corresponding matrix.

□

Finally, we are in the position to prove Theorem 5.0.1.

Proof of Theorem 5.0.1. The statement of the theorem follows by combining Theorem 4.2.2 with Proposition 5.2.2, Proposition 5.2.3 as well as Proposition 5.3.3 and Proposition 5.2.4.

□

Appendix

In this appendix we give a loose collection of, mostly well-known, results and formulas.

Let $X = (X_k)_{k\geq 0}$ be a Markov chain with finite (or countably infinite) state space $\mathcal{S}$. Furthermore, for a given $A \subset \mathcal{S}$ define $m_{x,A} \coloneqq \mathbb{E}_x \inf\{k \geq 0 |\ X_k \in A\}$ as the mean first entry time of A given a starting state x.

Proposition A.1. Fix $A \subset \mathcal{S}$ and denote by $\mathbf{P} = (p_{x,y})_{x,y\in\mathcal{S}}$ the one step transition matrix of X. Then, the vector $\mathbf{m}_A = (m_{x,A})_{x\in\mathcal{S}}$ is the minimal non-negative solution to the following linear system of equations:

$$m_{x,A} = 0 \quad \text{if } x \in A \quad \text{and} \quad m_{x,A} = 1 + \sum_{y\notin A} p_{x,y} m_{y,A} \quad \text{if } x \notin A.$$

A proof is given in [No], Theorem 1.3.5.

The next proposition is the famous Gershgorin Theorem that makes an a priori statement about the location of the eigenvalues of a given matrix $\mathbf{A}$.

Proposition A.2. (Gershgorin)

Let $\mathbf{A} = (a_{i,j})_{i,j=1}^K$ be an arbitrary $K \times K$-matrix with complex entries. Denote by G_i the i-th Gershgorin disk, i.e.

$$G_i \coloneqq \left\{ z \in \mathbb{C} |\ |z - a_{i,i}| \leq \sum_{j\neq i} |a_{i,j}| \right\}.$$

Then

$$\sigma(\mathbf{A}) \subset \bigcup_{i=1}^K G_i.$$

A proof for this statement can be found in [Va], Theorem 1.11.

Next we state the Hurwitz Theorem. We quote the formulation given in [RS], Theorem 1.3.8, where the reader can also find the proof. Recall the notation $B_R(a) \coloneqq \{z \in \mathbb{C} |\ |z-a| < R\}$.

Proposition A.3.(Hurwitz)

Let $G \subset \mathbb{C}$ be a region and let $(f_k)_{k\geq 1}$ be a sequence of analytic functions on G that converges to a non-zero function f uniformly on every compact subset of G. Then $z_0 \in G$ is a zero of f with multiplicity m if and only if there exists a neighborhood $H \subset G$ such that, in every ball $B_R(z_0) \subset H$, each function f_n whose index exceeds some bound $n_0 = n_0(R)$ has exactly m zeros, counted according to their multiplicities.

The next lemma is a generalization of the binomial formula and can be derived by a simple straightforward calculation.

Lemma A.4. *Fix numbers $c_1, \ldots, c_n, d_1, \ldots, d_n \in \mathbb{R}$. Then*

$$\prod_{i=1}^{n}(c_i + d_i) = \sum_{i=0}^{n} G_i$$

where

$$G_0 = \prod_{l=1}^{n} d_l \quad \textit{and} \quad G_i = \sum_{1\leq j_1<\ldots<j_i\leq n} \prod_{k\in\{j_1,\ldots,j_i\}} c_k \prod_{\substack{l\in\{1,\ldots,n\}\\ l\notin\{j_1,\ldots,j_i\}}} d_l.$$

List of symbols

$\sqcup$	disjoint union
$:=$	defined to be equal
$\stackrel{d}{=}$	equality in distribution
$\mathcal{B}(\mathbb{R})$	Borel sets of $\mathbb{R}$
$\mathbf{1}_A$	indicator function of an event A
$\mathbf{I}_K$	identity matrix of dimension $K \geq 1$
$\mathbf{0}_K$	the vector $(0, 0 \ldots, 0)^T$ of dimension $K \geq 1$
$\mathbf{1}_K$	the vector $(1, 1, \ldots, 1)^T$ of dimension $K \geq 1$
U	potential function
n	number of potential wells of U
Ω_i	i-th well of U
$\mathfrak{m}_i$	i-th local minimum of U
$\mathfrak{s}_i$	i-th local maximum of U
$\mathcal{M}$	set of all local minima of U, i.e. $\mathcal{M} = \{\mathfrak{m}_1, \ldots, \mathfrak{m}_n\}$
h	parameter of the time discretization
δ	parameter of the space discretization
R	boundary of the state space
S_N	set of all permutations $\boldsymbol{\pi} : \mathcal{S} \to \mathcal{S}$
S_n	set of all permutations $\boldsymbol{\sigma} : \{1, \ldots, n\} \to \{1, \ldots, n\}$

Bibliography

[App] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press (2004)

[Berg] H. Bergström, *On Some Expansions of Stable Distribution Functions*, Arkiv för Matematik, Volume 2, Issue 4, pp 375–378 (1952)

[BGK] A. Bovier,V. Gayrard, M. Klein, *Metastability in Reversible Diffusion Processes II: Precise Asymptotics for Small Eigenvalues*, Journal of the European Mathematical Society, Volume 7, Issue 1, pp 69–99 (2005)

[BEGK1] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, *Metastability in Reversible Diffusion Processes I: Sharp Asymptotics for Capacities and Exit Times*, Journal of the European Mathematical Society, Volume 6, Issue 4, pp 399-424 (2004)

[BEGK2] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, *Metastability and Low Lying Spectra in Reversible Markov Chains*, Communications in Mathematical Physics, Volume 228, Issue 2, pp 219–255 (2002)

[BuMa1] V.A. Buslov, K.A. Makarov, *Hierarchy of Time Scales in the Case of Weak Diffusion*, (Russian) Teoret. Mat. Fiz. 76, p. 219-230 (1988), translated in Theoretical and Mathematical Physics, Volume 76, Issue 2, pp 818–826 (1988)

[BuMa2] V.A. Buslov, K.A. Makarov, *Life Times and Lower Eigenvalues of an Operator of Small Diffusion*, (Russian) Mat. Zametki 51, pp 20-31 (1992), translated in Mathematical Notes, Volume 51, Issue 1, pp 14–21 (1992)

[Ca] M. K. Cameron, *Computing Freidlin's Cycles for the Overdamped Langevin Dynamics*, Journal of Statistical Physics, Volume 152, Issue 3, pp 493–518 (2013)

[CoDo] P. Coolen-Schrijner, E.A. van Doorn, *The Deviation Matrix of a Continuous-time Markov Chain*, Department of Applied Mathematics, University of Twente (2001), available at *http://doc.utwente.nl/65754/* (Last checked: 30.06.2014)

[Di1] P. D. Ditlevsen, *Observation of α-stable Noise and a Bistable Climate Potential in an Ice-core Record*, Geophysical Research Letters, Volume 26, Issue 10, pp 1441–1444 (1999)

[Di2] P. D. Ditlevsen, *Anomalous Jumping in a Double-well Potential*, Physical Review E, Volume 60, Issue 1, pp 172–179 (1999)

[Eck] M. Eckhoff, *Capacity and the Low Lying Spectrum in Attractive Markov Chains*, Dissertation, Universität Potsdam (2000)

[EtKu] S. N. Ethier, T. G. Kurtz, *Markov Processes — Characterization and Convergence*, John Wiley and Sons, Hoboken, New Jersey (1986)

[Fell] W. Feller, *An Introduction to Probabilty Theory and Its Applictaions*, Volume 2, Second Edition, Wiley (1971)

[Fr] D. Freedman, *Markov Chains*, Holden-Day Series in Probability and Statistics, Springer-Verlag New York (1983)

[FrGr] B. Fridstedt, L. Gray, *A Modern Approach to Probability Theorie*, Birkhäuser Boston (1997)

[FrWe] M.I. Freidlin, A.D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag New York (1984)

[GOV] A. Galves, E. Olivieri, M.E. Vares, *Metastability for a Class of Dynamical Systems Subject to Small Random Perturbations*, The Annals of Probability, Volume 15, Issue 4, pp 1288-1305 (1987)

[ImPa1] P. Imkeller, I. Pavlyukevich, *First Exit Times of SDEs Driven by Stable Lévy Processes*, Stochastic Processes and Their Applications, Volume 116, Issue 4, pp 611–642 (2006)

[ImPa2] P. Imkeller, I. Pavlyukevich, *Metastable Behavior of Small Noise Lévy-driven Diffusions*, ESAIM: Probability and Statistics, Volume 12, pp 412–437 (2008)

[Ja] N. Jacob, *Pseudo-differential Operators and Markov Processes*, Volume 3, Imperial College Press (2005)

[Ka] O. Kallenberg, *Foundations of Modern Probability*, Second Edition, Probability and Its Applications, Springer (2010)

[KaSh] I. Karatzas, S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Graduate Texts in Mathematics 113, Springer-Verlag New York (1991)

[KiNe] C. Kipnis and C.M. Newman, *The Metastable Behavior of Infrequently Observed, Weakly Random, One-dimensional Diffusion Processes*, SIAM Journal on Applied Mathematics, Volume 45, Issue 6, pp 972–982 (1985),

[Ko1] V.N. Kolokoltsov, *Symmetric Stable Laws and Stable-Like Jump Diffusions*, Proceedings of the London Mathematical Society, Volume 80, Issue 03, pp 725-768 (2000)

[Ko2] V.N. Kolokoltsov, *Semiclassical Analysis for Diffusions and Stochastic Processes*, Lecture Notes in Mathematics, Volume 1724, Springer-Verlag Berlin (2000)

[KoMa] V.N. Kolokoltsov, K.A. Makarov *Asymptotic Spectral Analysis of a Small Diffusion Operator and the Life Times of the Corresponding Diffusion Process*, Russian Journal of Mathematical Physics, Volume 4, Issue 3, pp 341–360 (1997)

[KeSn1] J. G. Kemeny, J. L. Snell, *Finite Markov Chains*, D. van Nostrand Company, Inc. (1960)

[KeSn2] J. G. Kemeny, J. L. Snell, *Finite Continuous Time Markov Chains*, Theory of Probability and Its Applications, Volume 6, Issue 1, pp 101–105 (1961)

[Li] T. M. Ligget, *Continuous Time Markov Processes — An Introduction*, American Mathematical Society (2010)

[No] J. R. Norris, *Markov Chains*, Cambridge University Press (1998)

[Ok] B. Oksendal, *Stochastic Differential Equations — An Introduction with Applications*, 6th Edition, Corrected 3th printing, Springer-Verlag Berlin (2006)

[Pr] P. Protter, *Stochastic Integration and Differential Equations*, Springer-Verlag Berlin (1990)

[RS] Q. I. Rahman, G. Schmeisser, *Analytic Theory of Polynomials*, Oxford University Press (2002)

[Sa] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press (1999)

[SaTa] G. Samorodnitsky, M.S. Taqqu, *Stable Non-Gaussian Random Processes*, Stochastic Modeling, New York: Chapman and Hall (1994)

[Sys] R. Syski, *Ergodic Potentials*, Stochastic Processes and Their Applications, Volume 7, Issue 3, pp 311-336 (1978)

[Va] R.S. Varga, *Matrix Iterative Analysis*, Second Edition, Springer (2010)

[Whi] W. Whitt, *Asymptotic Formulas for Markov Processes with Applications to Simulation*, Operations Research, Volume 40, Issue 2, pp 279-291 (1992)

[Xia] A. Xia, *Weak Convergence of Jump Processes*, Lecture Notes in Mathematics, Volume 1526. Séminaire de Probabilités XXVI, pp 32-46 (1992)

[Zo] V. M. Zolotarev, *One-dimensional Stable Distributions*, Translation of Mathematical Monographs 65, American Mathematical Society, Providence RI (1986)

Ehrenwörtliche Erklärung

Hiermit erkläre ich,

- dass mir die Promotionsordnung der Fakultät bekannt ist,

- dass ich die Dissertation selbst angefertigt habe, keine Textabschnitte oder Ergebnisse eines Dritten oder eigenen Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönliche Mitteilungen und Quellen in meiner Arbeit angegeben habe,

- dass ich die Hilfe eines Promotionsberaters nicht in Anspruch genommen habe und daß Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,

- dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts haben mich folgende Personen unterstützt:

.....

.....

Ich habe die gleiche, eine in wesentlichen Teilen ähnliche bzw. eine andere Abhandlung* bereits bei einer anderen Hochschule als Dissertation eingereicht: Ja / Nein*.
(* Zutreffendes unterstreichen)

Wenn Ja, Name der Hochschule: .....

Ergebnis: .....

Jena, den .....

Unterschrift